

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 ; Search time 84.8957 Seconds
(without alignments)
108.037 Million cell updates/sec

Title: US-09-932-322-1
Perfect score: 44
Sequence: 1 XXXCXXPTGCGXXX 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	81.8	192	2	Q98A99_RHILIO
2	36	81.8	360	2	Q86AK7_DICDI
3	36	81.8	546	2	Q8KJ72_RHILIO
4	36	81.8	597	2	Q5AK44_DICDI
5	36	81.8	601	2	Q7M4J3_DICDI
6	36	81.8	608	2	Q627A0_CABBR
7	36	81.8	1710	2	Q54Y62_DICDI
8	35	79.5	148	2	Q95KD2_MACFA
9	35	79.5	148	2	Q9N018_MACFA
10	35	79.5	370	2	Q5SBT6_DICDI
11	35	79.5	370	2	Q5SBX0_DICDI
12	35	79.5	466	2	Q5KKP3_CRYNE
13	35	79.5	571	2	Q9AOP5_PSEBE
14	35	79.5	997	2	Q4SBN7_TETNG
15	35	79.5	1018	2	Q4SP52_TETNG
16	35	79.5	1839	2	Q4S795_TETNG
17	34	77.3	113	2	Q726J7_HUMAN
18	34	77.3	145	1	PA2X_NOTSC
19	34	77.3	167	1	Q7U7V9_SYNPX
20	34	77.3	296	2	Q54YF0_DICDI
21	34	77.3	312	2	Q9SUV3_ARATH
22	34	77.3	380	2	Q5AN14_DICDI
23	34	77.3	465	2	Q6TVH9_9POXV
24	34	77.3	487	2	Q4MWC1_ASPPU
25	34	77.3	504	2	Q52AZ8_ORYSA
26	34	77.3	507	2	Q5AN64_DICDI
27	34	77.3	670	2	Q4SPD9_TETNG
28	34	77.3	870	2	Q5B3W6_EMBNI
29	34	77.3	932	2	Q54N02_DICDI
30	34	77.3	998	2	Q869K4_DICDI
31	34	77.3	999	2	Q55C92_DICDI

32	34	77.3	1147	2	Q57XW5_PTEYP
33	33	77.3	2217	2	Q6Y625_YVIRU
34	33	75.0	77	2	Q4IR98_GIBRE
35	33	75.0	269	2	Q67W06_ORYSA
36	33	75.0	330	2	Q5NVV4_9ARCH
37	33	75.0	464	2	Q7XV21_ORYSA
38	33	75.0	1058	2	Q9LEB5_ARATH
39	33	75.0	1468	2	Q80TF6_MOUSE
40	32	72.7	64	2	Q92U25_RHIME
41	32	72.7	123	2	Q926M4_LISIN
42	32	72.7	222	2	Q5RN00_ANGOA
43	32	72.7	231	2	Q5TMZ9_ANGOA
44	32	72.7	520	2	Q628K1_ORYSA
45	32	72.7	1152	2	Q9FI26_ARATH
46	31	70.5	162	2	Q9WU17_MESAU
47	31	70.5	166	2	Q8T415_DROME
48	31	70.5	183	2	Q4H6G8_9PEIO
49	31	70.5	229	2	Q4S7H4_TETNG
50	31	70.5	232	2	Q93KV9_STRVR
51	31	70.5	359	2	Q6VMH4_STEAM
52	31	70.5	361	2	Q8ECT9_SHEON
53	31	70.5	482	2	Q9LZB8_ARATH
54	31	70.5	615	2	Q22866_CAEEL
55	31	70.5	719	2	Q54I92_DICDI
56	30	68.2	66	2	Q7N375_PHOLL
57	30	68.2	78	2	P90569_PLABE
58	30	68.2	95	2	P77130_ECOLI
59	30	68.2	100	1	CHA3_BOMMO
60	30	68.2	114	2	Q8WQ78_CAEEL
61	30	68.2	119	2	Q4UUP1_XANCP
62	30	68.2	119	2	Q8P958_XANCP
63	30	68.2	129	1	CHA1_BOMMO
64	30	68.2	132	1	CHA2_BOMMO
65	30	68.2	133	2	Q9LIW4_ORYSA
66	30	68.2	150	2	Q6D216_ERMCT
67	30	68.2	151	2	Q8ZD76_YERPE
68	30	68.2	151	2	Q66707_YERPS
69	30	68.2	153	2	Q8KQC9_YERSIN
70	30	68.2	157	2	Q614N5_CABBR
71	30	68.2	171	2	Q9ULP0_RAT
72	30	68.2	175	2	Q6IJNO_DROME
73	30	68.2	197	2	Q4LG50_9BORK
74	30	68.2	197	2	Q4U1V1_9SMIG
75	30	68.2	199	2	Q7RUV7_NEUCR
76	30	68.2	200	1	WIN1_SOLRV
77	30	68.2	201	2	Q03954_LYCBS
78	30	68.2	203	2	Q7SQH1_CAPCH
79	30	68.2	208	2	Q41230_TOBAC
80	30	68.2	210	2	Q9SEEM3_CAPAN
81	30	68.2	211	1	WIN2_SOLRV
82	30	68.2	211	2	Q41231_TOBAC
83	30	68.2	222	1	SP12B_VACCC
84	30	68.2	222	2	Q6J350_9POXV
85	30	68.2	222	2	Q9UPF3_VACCT
86	30	68.2	224	2	Q9HP93_HAL9A
87	30	68.2	232	2	Q9BBE9_PIG
88	30	68.2	239	2	Q25653_PLABE
89	30	68.2	240	2	Q59587_9ACTO
90	30	68.2	246	2	Q5RNY6_ANGOA
91	30	68.2	263	2	Q58827_ECTHA
92	30	68.2	290	2	Q4G0U2_HUMAN
93	30	68.2	299	2	Q5SM71_CRYNE
94	30	68.2	299	2	Q5KB57_CRYLOCC
95	30	68.2	329	2	Q6CL1K_YARLI
96	30	68.2	336	2	Q9KDK4_BACCH
97	30	68.2	345	1	SP12_RABPU
98	30	68.2	345	1	SP12_VACCV
99	30	68.2	345	1	Q72749_COMXP
100	30	68.2	347	2	Q4KSK4_9NUCL
101	30	68.2	359	2	Q4PLCQ_SCHMA
102	30	68.2	397	2	Q7NKS5_GLOVI
103	30	68.2	406	2	Q21712_CABEL
104	30	68.2	423	2	Q4U1U8_9SMEG

Q57XW5	trypanosoma
Q6Y625	iasia virus
Q4IR98	gibberella
Q67W06	oryza sativ
Q5NVV4	uncultured
Q7XV21	oryza sativ
Q9LEB5	arabidopsis
Q80TF6	mus musculus
Q92U25	ribobium m
Q926M4	listeria in
Q5RN00	anopheles g
Q5TMZ9	anopheles g
Q628K1	oryza sativ
Q9FI26	arabidopsis
Q9WU17	mesocricetu
Q8T415	drosophila
Q4H6G8	deinococcus
Q4S7H4	tetradon n
Q93KV9	streptomyce
Q6VMH4	streptomyce
Q8ECT9	shewanella
Q9LZB8	arabidopsis
Q22866	caenorhabdi
Q54I92	dictyosteli
Q7N375	photornabdu
P90569	plasmodium
P77130	escherichia
P08929	bomblyx mori
Q8WQ78	caenorhabdi
Q4UUP1	xanthomonas
Q8P958	xanthomonas
P08826	bomblyx mori
P08825	bomblyx mori
Q9LIW4	oryza sativ
Q6IJNO	erynia car
Q8ZD76	yersinia ps
Q66707	yersinia ps
Q8KQC9	yersinia en
Q614N5	caenorhabdi
Q9ULP0	rattus norv
Q6IJNO	drosophila
Q4LG50	burkholderi
Q4U1V1	pungitius p
Q7RUV7	neutrospora
P09761	solanum tub
Q03954	lycopersico
Q7SQH1	capsicum ch
Q41230	nicotiana t
Q9SEEM3	capsicum an
P09762	solanum tub
Q41231	nicotiana t
P20842	vaccinia vi
Q6J350	vaccinia vi
Q9UPF3	vaccinia vi
Q9HP93	halobacteri
Q9BBE9	sw sacrofa
Q25653	plasmodium
Q59587	streptomyce
Q5RNY6	anopheles g
Q58827	ectothiorio
Q4G0U2	homo sapien
Q5SM71	cryptococcu
Q5KB57	cryptococcu
Q6CL1K	yarrowia li
Q9KDK4	bacillus ha
P42826	rabidipox v
P15059	vaccinia vi
Q72749	compos viru
Q4KSK4	chrysodeixi
Q4PLCQ	schistosoma
Q7NKS5	gloeobacter
Q21712	caenorhabdi
Q4U1U8	pungitius p

105	30	68.2	426	2	04LXZ5_9BURK	04LXZ5_burkholderi	178	29	65.9	162	2	06Z1J3_STRAW	06Z1J3_streptomyce
106	30	68.2	430	2	021713_CABEL	021713_caenorhabdi	179	29	65.9	163	2	061832_CABEL	061832_caenorhabdi
107	30	68.2	442	2	084170_9POXV	084170_orf_virus.	180	29	65.9	164	2	061H39_CAEBR	061H39_caenorhabdi
108	30	68.2	452	2	08TWY8_METKA	08TWY8_methanopyru	181	29	65.9	171	2	06RR40_ARATH	06RR40_arabidopsi
109	30	68.2	452	2	06TWY1_9POXV	06TWY1_orf_virus.	182	29	65.9	195	2	095QZ5_CABEL	095QZ5_caenorhabdi
110	30	68.2	453	2	06TW91_9POXV	06TW91_orf_virus.	183	29	65.9	205	2	04FZ56_PENMO	04FZ56_pennaeus non
111	30	68.2	479	2	07RYP8_NEUCR	07RYP8_neuropospora	184	29	65.9	201	2	061K16_DROME	061K16_drosophila
112	30	68.2	493	2	09ATG8_CHLRE	09ATG8_chlamydomon	185	29	65.9	207	2	083NCG_TROM8	083NCG_tropheryma
113	30	68.2	527	1	MAS13_CABEL	020191_caenorhabdi	186	29	65.9	211	2	08YKJ3_MYCTU	08YKJ3_mycobacteri
114	30	68.2	551	1	AKS_HEMPU	P14000_hemiteictric	187	29	65.9	219	2	000048_AJDEE	000048_ajelilomyces
115	30	68.2	551	2	AKS_HEMPU	P14000_hemiteictric	188	29	65.9	226	2	0673H2_PENVA	0673H2_pennaeus van
116	30	68.2	558	2	052384_HEMPU	025384_hemiteictric	189	29	65.9	226	2	04FZ57_LITSE	04FZ57_litopenaeus
117	30	68.2	572	1	CECR6_MOUSE	0522GF0_magrapothe	190	29	65.9	226	2	07UL27_RHOBA	07UL27_rhodopiirell
118	30	68.2	574	2	076632_NEUCR	099XK70_mus_musculu	191	29	65.9	228	2	04SRJ6_TETNG	04SRJ6_tetradodon n
119	30	68.2	582	2	06MPL4_NEUCR	076632_neuropospora	192	29	65.9	231	2	051KW9_MAGGR	051KW9_magrapothe
120	30	68.2	619	2	04PJU7_USITMA	06MPL4_neutrospora	193	29	65.9	237	2	054V93_DICDI	054V93_dictyosteli
121	30	68.2	635	1	GBRT_MOUSE	04PJU7_ustiliago ma	194	29	65.9	245	2	09ZUV1_ARATH	09ZUV1_arabidopsi
122	30	68.2	635	2	0912M7_RAT	091EM7_rattus norv	195	29	65.9	254	2	06L0G4_PICVO	06L0G4_picrophilus
123	30	68.2	642	2	0607E9_METCA	0911F1_mus_musculu	196	29	65.9	262	2	07NJ98_GLOVI	07NJ98_gloeobacter
124	30	68.2	650	2	04SSS2_TETNG	0607E9_methylococc	197	29	65.9	263	2	051S12_MAGGR	051S12_magrapothe
125	30	68.2	699	2	08NB41_HUMAN	04SSS2_tetradodon n	198	29	65.9	271	2	06WVJ7_PENMO	06WVJ7_pennaeus mon
126	30	68.2	707	2	08COW9_MOUSE	08NB41_homo sapien	199	29	65.9	274	2	07UI56_RHOBA	07UI56_rhodopiirell
127	30	68.2	708	2	07YSJ4_DICDI	08COW9_mus_musculu	200	29	65.9	276	2	093X97_ANTHI	093X97_antirrhinum
128	30	68.2	715	2	094494_DICDI	07Y6J4_dictyosteli	201	29	65.9	276	2	04T6Q2_TETNG	04T6Q2_tetradodon n
129	30	68.2	744	2	04TV95_TETNG	094494_dictyosteli	202	29	65.9	281	2	07UNK9_RHOBA	07UNK9_rhodopiirell
130	30	68.2	746	2	054CH8_DICDI	04TV95_tetradodon n	203	29	65.9	283	2	05VOKO_ORYSA	05VOKO_oryza sativ
131	30	68.2	764	2	09HAP8_HUMAN	054CH8_dictyosteli	204	29	65.9	293	1	CHIE_BEYU	06F7P8_beta vulgar
132	30	68.2	785	2	08TKK7_METAC	09HAP8_homo sapien	205	29	65.9	294	2	06F7P8_ACIAD	06F7P8_acinetobact
133	30	68.2	789	2	09H8C9_HUMAN	08TKK7_methanoscac	206	29	65.9	293	2	061H14_CAEBR	061H14_caenorhabdi
134	30	68.2	834	2	05JHL1_PYRO	09H8C9_homo sapien	207	29	65.9	326	2	074GE6_GEOSL	074GE6_geobacter b
135	30	68.2	884	1	ANDR_EDLFC	05JHL1_pyrococcus	208	29	65.9	330	2	07M725_WOLSU	07M725_wolfinella s
136	30	68.2	896	1	ANDR_PIG	097776_eulemur ful	209	29	65.9	333	2	0980J4_SULISO	0980J4_sulfolobus
137	30	68.2	966	2	05ACJ1_DICDI	098J17_sus scrofa	210	29	65.9	335	2	057UB4_GTRYP	057UB4_trypanosoma
138	30	68.2	1157	2	051UM6_MAGGR	05ACJ1_dictyosteli	211	29	65.9	336	2	06ZVW7_HUMAN	06ZVW7_homo sapien
139	30	68.2	1245	2	0962K7_PLAVI	051UM6_magrapothe	212	29	65.9	345	2	061G18_DROME	061G18_drosophila
140	30	68.2	1367	2	04PHL3_USITMA	0962K7_plamodiom	213	29	65.9	348	2	09X7M9_STRCO	09X7M9_streptomyce
141	30	68.2	1375	2	054YB9_DICDI	04PHL3_ustiliago ma	214	29	65.9	350	2	060768_9BURK	060768_burkholderi
142	30	68.2	1556	2	04Z273_PLABE	054YB9_dictyosteli	215	29	65.9	350	2	082MH4_STEAM	082MH4_streptomyc
143	30	68.2	1721	2	0614N6_CABER	04Z273_plamodiom	216	29	65.9	352	2	05FW56_MOUSE	05FW56_mus_musculu
144	30	68.2	1722	2	019350_CABEL	0614N6_caenorhabdi	217	29	65.9	351	2	0500Q5_PSESY	0500Q5_pseudomonas
145	30	68.2	1787	2	025645_PLABE	019350_caenorhabdi	218	29	65.9	351	2	088B73_PSESM	088B73_pseudomonas
146	30	68.2	1905	2	0906D4_PLABE	025645_plamodiom	219	29	65.9	353	2	08VPM3_9MICC	08VPM3_micrococcus
147	30	68.2	1972	2	04Q6D0_LEIMA	0906D4_leishmania	220	29	65.9	356	2	04FVU6_LEIMA	04FVU6_leishmania
148	30	68.2	1972	2	07PUY1_ANOGA	04Q6D0_leishmania	221	29	65.9	357	2	052B09_MAGGR	052B09_magrapothe
149	30	68.2	2161	2	091724_9VIRU	07PUY1_anopheles g	222	29	65.9	366	2	041ZRS_AZOVI	041ZRS_azobacter
150	30	68.2	2161	2	052W13_9VIRU	091724_rupestrictis s	223	29	65.9	369	2	04KKM4_PSEFF	04KKM4_pseudomonas
151	30	68.2	2161	2	052W18_9VIRU	052W13_rupestrictis s	224	29	65.9	375	2	091766_PSEAB	091766_pseudomonas
152	30	68.2	23015	2	081O18_DROME	052W18_rupestrictis s	225	29	65.9	379	2	004326_PICAN	004326_pichia angu
153	30	68.2	23015	2	091W92_TOBAC	081O18_drosophila	226	29	65.9	383	2	05P3Z8_AZOSE	05P3Z8_azarcus sp
154	29	65.9	45	2	Y940_TEBPA	091W92_nicotiana t	227	29	65.9	385	2	08DLR1_SYNEK	08DLR1_synechococc
155	29	65.9	46	1	Y940_TEBPA	083910_treponema p	228	29	65.9	387	2	088RN6_PSEPK	088RN6_pseudomonas
156	29	65.9	76	2	04SYE8_TETNG	048Y68_tetradodon n	229	29	65.9	394	1	K1H4_HUMAN	06G7M2_bifidobacte
157	29	65.9	94	2	075079_HUMAN	04SYE8_tetradodon n	230	29	65.9	395	2	08G7M2_BIFLO	08G7M2_bifidobacte
158	29	65.9	95	2	05TA77_HUMAN	075079_homo sapien	231	29	65.9	404	2	09PULO_BRABE	09PULO_brechydantio
159	29	65.9	116	2	06RE47_CAPHI	05TA77_homo sapien	232	29	65.9	407	2	05B3N5_EMENI	05B3N5_aspergillus
160	29	65.9	118	2	053NF3_ORYSA	06RE47_cappi hircu	233	29	65.9	415	2	05KC03_CRYNE	05KC03_cryptococcu
161	29	65.9	119	1	CHCA1_BOOMO	053NF3_oryza sativ	234	29	65.9	429	2	08MRX5_DROME	08MRX5_drosophila
162	29	65.9	119	1	CHCA2_BOOMO	P13531_bombyx mori	235	29	65.9	429	2	09VNP4_DROME	09VNP4_drosophila
163	29	65.9	119	1	CHCA3_BOOMO	017212_bombyx mori	236	29	65.9	435	2	04S0U5_TETNG	04S0U5_tetradodon n
164	29	65.9	121	1	CHCA5_BOOMO	017212_bombyx mori	237	29	65.9	436	2	081UT8_HUMAN	081UT8_homo sapien
165	29	65.9	123	2	05N7A4_ORYSA	P08829_bombyx mori	238	29	65.9	444	2	051KF5_MAGGR	051KF5_magrapothe
166	29	65.9	126	2	058HG2_9MXI	05N7A4_oryza sativ	239	29	65.9	446	2	0965K4_ARATH	0965K4_arabidopsi
167	29	65.9	129	2	06ZSH8_HUMAN	058HG2_polyascus p	240	29	65.9	464	2	0866K1_GTRYP	0866K1_trypanosoma
168	29	65.9	129	2	06ZSH8_HUMAN	06ZSH8_homo sapien	241	29	65.9	469	2	08SMW9_DROME	08SMW9_drosophila
169	29	65.9	131	2	05OLU9_ORYSA	05QJ19_oryza sativ	242	29	65.9	472	2	06QA68_9TRYP	06QA68_trypanosoma
170	29	65.9	133	2	08MO77_CABEL	08Mq77_caenorhabdi	243	29	65.9	475	2	04Q4Q0_LEIMA	04Q4Q0_leishmania
171	29	65.9	147	2	09PY8_XYLLA	09PY8_xyella fas	244	29	65.9	477	2	08H3W6_ORYSA	08H3W6_oryza sativ
172	29	65.9	151	2	04QCK6_LEIMA	04QCK6_leishmania	245	29	65.9	481	2	058DB9_BOVIN	058DB9_bos taurus
173	29	65.9	151	2	07YS88_PIG	07YS88_sus scrofa	246	29	65.9	484	2	0529S0_MAGGR	0529S0_magrapothe
174	29	65.9	151	2	0867B8_PIG	0867B8_sus scrofa	247	29	65.9	489	2	05WRM0_CABEL	05WRM0_caenorhabdi
175	29	65.9	153	2	05U614_HUMAN	05U614_homo sapien	248	29	65.9	496	2	09RGS1_BACCI	09RGS1_bacillus ci
176	29	65.9	161	2	061HS7_CAEBR	061HS7_caenorhabdi	249	29	65.9	504	2	054FDB_DICDI	054FDB_dictyosteli
177	29	65.9	161	2	061HS7_CAEBR	061HS7_caenorhabdi	250	29	65.9	504	2	054FDB_DICDI	054FDB_dictyosteli

251	29	65.9	506	2	Q600H2_CABER	Q60h2 caenorhabdi	324	28	63.6	74	1	MS94B_DROME	Q01643 dtrosophila
252	29	65.9	507	2	Q9XU79_CABEL	Q9XU79 caenorhabdi	325	28	63.6	79	2	Q6K3A1_ORYSA	Q6K3a1 oryza sativ
253	29	65.9	515	1	APX1_CABEL	P41990 caenorhabdi	326	28	63.6	74	2	Q96U19_AJECA	Q96U19 ajeelomyces
254	29	65.9	515	2	Q4T2J6_TETNG	Q4c7j6 tetraodon n	327	28	63.6	86	2	Q8UMD7_9H1V1	Q8UMD7 human immun
255	29	65.9	541	1	DCR1B_MOUSE	Q8c7j6 mus musculu	328	28	63.6	91	2	Q9KCR3_9CUCU	Q9KCR3 drosophila
256	29	65.9	547	2	Q7UVV3_RHOBA	Q7UVV3 rhodoptreil	329	28	63.6	91	2	Q18928_MACRA	Q18928 macera radi
257	29	65.9	551	2	Q9DIC6_9RHAB	Q9dic6 letuce nec	330	28	63.6	95	2	Q8URK1_AGR75	Q8URK1 agrobacteri
258	29	65.9	574	2	Q9BMT7_DROME	Q9bmt7 dtrosophila	331	28	63.6	96	2	Q96419_9CUCU	Q96419 drosophila
259	29	65.9	574	2	Q9VJ22_DROME	Q9VJ22 dtrosophila	332	28	63.6	96	2	Q5LAC3_BACFN	Q5LAC3 bacteroides
260	29	65.9	574	2	Q6GWC3_BBARB	Q6gmj3 brachydantio	333	28	63.6	98	2	Q9WC06_9CUCU	Q9WC06 drosophila
261	29	65.9	575	2	Q7QNV8_GIALA	Q7qnv8 giardia lam	334	28	63.6	101	2	Q4QXB8_9H1V1	Q4QXB8 human immun
262	29	65.9	578	2	Q57U16_9TRYP	Q57u16 trypanosoma	335	28	63.6	103	2	Q49320_ARYTH	Q49320 arabidopsis
263	29	65.9	575	2	Q4T917_TETNG	Q4t917 tetraodon n	336	28	63.6	105	2	Q17187_BOMBO	Q17187 bombyx mori
264	29	65.9	612	2	Q9KZU9_STRCO	Q9KZU9 streptomyces	337	28	63.6	106	2	Q9G557_9HYME	Q9G557 xylocopa ol
265	29	65.9	614	2	Q4RLC5_TETNG	Q4RLC5 tetraodon n	338	28	63.6	107	2	Q8RKW5_PRORE	Q8RKW5 providencia
266	29	65.9	630	2	Q6TNG4_PICPA	Q6Tng4 picchia fari	339	28	63.6	108	2	Q46351_9CUCU	Q46351 drosophila
267	29	65.9	680	2	Q72MM5_DESVH	Q72mm5 desulfovibr	340	28	63.6	108	2	Q46352_9CUCU	Q46352 drosophila
268	29	65.9	685	1	DL4_HUMAN	Q9n161 homo sapien	341	28	63.6	108	2	Q5LAC4_BACFN	Q5LAC4 bacteroides
269	29	65.9	686	1	DL4_MOUSE	Q9j171 mus musculu	342	28	63.6	108	2	Q64Q05_BACFR	Q64Q05 bacteroides
270	29	65.9	686	2	Q9DBU9_MOUSE	Q9dbu9 mus musculu	343	28	63.6	109	2	Q46346_9CUCU	Q46346 drosophila
271	29	65.9	728	2	Q54DV5_DICDI	Q54dv5 dictyosteli	344	28	63.6	110	2	Q15739_HUMAN	Q15739 homo sapien
272	29	65.9	733	2	Q54P69_DICDI	Q54p69 dictyosteli	345	28	63.6	110	2	Q15739_HUMAN	Q15739 homo sapien
273	29	65.9	748	2	Q86HR8_DICDI	Q86hr8 dictyosteli	346	28	63.6	115	1	MERT_SHEPU	Q54462 shewanella
274	29	65.9	764	2	Q97343_SUEBA	Q97343 suberites d	347	28	63.6	120	2	Q86BW2_ASCSS	Q86bw2 ascidia syd
275	29	65.9	830	2	Q6INMO_XENILA	Q6inmo xenopus lae	348	28	63.6	121	2	Q9NCR1_9CUCU	Q9NCR1 bradyrhizob
276	29	65.9	843	1	CO7_HUMAN	P10643 homo sapien	349	28	63.6	121	2	Q9RMS0_BRAJA	Q9RMS0 xylocopa ir
277	29	65.9	843	1	CO7_PIG	Q9rtq3 sus scrofa	350	28	63.6	122	2	Q9NCR9_9CUCU	Q9NCR9 drosophila
278	29	65.9	843	2	Q5R4D0_PONPY	Q5r4d0 pongo pygma	351	28	63.6	123	2	Q9NCR9_9CUCU	Q9NCR9 drosophila
279	29	65.9	919	2	Q61V24_CABER	Q61v24 caenorhabdi	352	28	63.6	129	2	Q9NCR9_9CUCU	Q9NCR9 drosophila
280	29	65.9	934	2	Q7XVG0_ORYSA	Q7xvg0 oryza sativ	353	28	63.6	131	2	Q4T174_TETNG	Q4T174 tetraodon n
281	29	65.9	966	2	Q22378_CABEL	Q22378 caenorhabdi	354	28	63.6	132	2	Q7XJVO_ORYSA	Q7XJVO oryza sativ
282	29	65.9	1005	2	Q5ZM51_CHICK	Q5zm51 gallus gall	355	28	63.6	136	2	Q9G567_9HYME	Q9G567 xylocopa tr
283	29	65.9	1029	2	Q9VW85_DROME	Q9vw85 dtrosophila	356	28	63.6	139	2	Q9G573_9HYME	Q9G573 xylocopa tr
284	29	65.9	1053	2	Q69ZC3_MOUSE	Q69zc3 mus musculu	357	28	63.6	141	2	Q8AW44_BBARB	Q8AW44 brachydantio
285	29	65.9	1070	2	Q5AKU3_EMENT	Q5akj3 aspergillus	358	28	63.6	143	2	Q9G566_XYLV1	Q9G566 xylocopa vi
286	29	65.9	1085	2	Q4SEL7_TETNG	Q4sel7 tetraodon n	359	28	63.6	144	2	Q5N2D3_SYNP6	Q5n2d3 synechococ
287	29	65.9	1094	2	Q7SD86_NEUCR	Q7sd86 neuropepora	360	28	63.6	146	2	Q4WG18_ASEPT	Q4WG18 aspergillus
288	29	65.9	1106	2	P79726_BBARB	P79726 brachydantio	361	28	63.6	146	2	Q5VTS2_HUMAN	Q5VTS2 homo sapien
289	29	65.9	1142	2	Q615H8_ORYSA	Q615h8 oryza sativ	362	28	63.6	146	2	Q4YMU1_PLABE	Q4YMU1 plasmodium
290	29	65.9	1183	2	Q7SDP1_NEUCR	Q7sdp1 neuropepora	363	28	63.6	146	2	Q8BSG4_MOUSE	Q8BSG4 mus musculu
291	29	65.9	1224	2	Q4SEB5_TETNG	Q4seb5 tetraodon n	364	28	63.6	147	2	Q9G571_9HYME	Q9G571 xylocopa vi
292	29	65.9	1335	2	Q9FME7_ARYTH	Q9fme7 arabidopsis	365	28	63.6	147	2	Q5TVP3_ANOGA	Q5TVP3 anopheles g
293	29	65.9	1357	2	Q5MD89_BBARB	Q5md89 brachydantio	366	28	63.6	149	2	Q8AVJ9_BBARB	Q8AVJ9 brachydantio
294	29	65.9	1357	2	Q5GIT2_BBARB	Q5git2 brachydantio	367	28	63.6	152	2	Q8FYJ8_BRURU	Q8FYJ8 bruceella su
295	29	65.9	1387	2	Q4GIQ3_TETNG	Q4giq3 tetraodon n	368	28	63.6	153	2	Q9G564_9HYME	Q9G564 xylocopa su
296	29	65.9	1417	2	Q4QCZ2_LEIMA	Q4qcz2 leishmania	369	28	63.6	153	2	Q9G570_9HYME	Q9G570 xylocopa bo
297	29	65.9	1420	2	Q4SCY1_TETNG	Q4scy1 tetraodon n	370	28	63.6	153	2	Q9G572_9HYME	Q9G572 xylocopa au
298	29	65.9	1423	1	Q5SW75_MOUSE	Q5sw75 mus musculu	371	28	63.6	154	2	Q9YAZ4_AERYP	Q9YAZ4 aetopyrum p
299	29	65.9	1427	1	MES4_DROME	Q8mw36 dtrosophila	372	28	63.6	156	2	Q6V262_9EURY	Q6V262 uncultured
300	29	65.9	1564	2	Q8GMD7_ECHMU	Q8gmd7 echinococcu	373	28	63.6	158	2	Q7MT99_PORDI	Q7MT99 pophyromon
301	29	65.9	1637	2	Q9XSU8_BOVIN	Q9xsu8 bos taurus	374	28	63.6	162	2	Q9BLH6_APLUK	Q9BLH6 aplysia kur
302	29	65.9	1743	2	Q9XWXS_CABEL	Q9xwx5 caenorhabdi	375	28	63.6	163	2	Q4SEL6_TETNG	Q4SEL6 tetraodon n
303	29	65.9	1748	2	Q61TPI_CABER	Q61tpi caenorhabdi	376	28	63.6	164	2	Q5DY73_VIBP1	Q5DY73 vibrio fisc
304	29	65.9	1755	2	Q55BY2_DICDI	Q55by2 dictyosteli	377	28	63.6	165	2	Q26742_9TRYP	Q26742 trypanosoma
305	29	65.9	1843	2	Q9GQ52_LEITR	Q9gq52 leishmania	378	28	63.6	168	2	Q5NXd6_AZORSE	Q5NXd6 azaracus sp
306	29	65.9	1917	2	Q51SW3_ENTHI	Q51sw3 entamoeba h	379	28	63.6	173	2	Q5IFP5_HORSE	Q5IFP5 equus cabal
307	29	65.9	2117	2	Q4OIS6_LEIMA	Q4ois6 leishmania	380	28	63.6	176	2	Q879W2_XYLF1	Q879W2 xyella fas
308	29	65.9	2146	2	Q4T7A2_TETNG	Q4t7a2 tetraodon n	381	28	63.6	176	2	Q891S1_BRAJA	Q891S1 bradyrhizob
309	29	65.9	2183	2	Q64962_9VIRU	Q64962 apple stem	382	28	63.6	176	2	Q6J340_9PROX	Q6J340 vaccinia vi
310	29	65.9	2185	2	Q9E948_9VIRU	Q9e948 apple stem	383	28	63.6	177	2	Q6AS16_ORYSA	Q6AS16 oryza sativ
311	29	65.9	2938	2	Q4IR68_GIBZE	Q4ir68 gibberella	384	28	63.6	178	2	P74672_SYNY3	P74672 synechocyst
312	29	65.9	2938	2	Q4G444_XENILA	Q4g444 xenopus lae	385	28	63.6	194	2	Q90YC9_CHICK	Q90YC9 gallus gall
313	29	65.9	3695	1	LAMAS_HUMAN	Q15J30 homo sapien	386	28	63.6	196	1	ZCSL3_MOUSE	Q912F0 mus musculu
314	29	65.9	3695	2	Q8TDF6_HUMAN	Q8tdf6 homo sapien	387	28	63.6	198	2	Q9RCU7_STRCO	Q9RCU7 streptomyces
315	29	65.9	3718	1	LAMAS_MOUSE	Q61001 mus musculu	388	28	63.6	201	2	Q9ANB9_BRAJA	Q9ANB9 bradyrhizob
316	29	65.9	4998	2	Q8CG65_MOUSE	Q8cg65 mus musculu	389	28	63.6	202	2	Q8H717_ORYSA	Q8H717 oryza sativ
317	29	65.9	5141	2	Q700K0_RAP	Q700k0 rattus norv	390	28	63.6	202	2	Q8G796_ORYSA	Q8G796 oryza sativ
318	29	65.9	5146	2	Q8SPM4_BOVIN	Q8spm4 bos taurus	391	28	63.6	203	2	Q27654_9NEOB	Q27654 extoxocrania
319	28	63.6	25	1	LRP3_CONPU	P60245 conus purpu	392	28	63.6	207	1	PLAB2_ARCFU	Q52J07 magnaorblo
320	28	63.6	25	1	LRP3_CONPU	P83629 felis silve	393	28	63.6	207	2	Q52G10_MAGGR	Q52G10 magnacorblo
321	28	63.6	31	2	Q4YAB1_PLABE	Q4yab1 plasmodium	394	28	63.6	207	2	Q9ZM78_HELPU	Q9ZM78 helicobacte
322	28	63.6	55	2	Q7UXB5_RHOBA	Q7uxb5 rhodoptreil	395	28	63.6	209	2	P4146_TRILE	P4146 trichosanth
323	28	63.6	59	2	Q6ADE4_LEITXX	Q6ade4 leifsonia x	396	28	63.6	210	2	Q7XJ92_WHEAT	Q7XJ92 triticum ae

397	28	63.6	215	2	051P06_MAGGR	051P06_magnaporthe	470	28	63.6	450	1	LIPP_PIG	P00591_sus_scrofa
398	28	63.6	215	2	04S216_TERNG	04S216_tetradodon n	471	28	63.6	450	2	06CAT3_VARI1	06CAT3_yarrowia li
399	28	63.6	218	2	09T7F3_CABEL	09T7F3_caenorhabdi	472	28	63.6	462	2	08MOC6_CABEL	08MOC6_caenorhabdi
400	28	63.6	234	2	07OGY2_ANOGA	07GY2_anopheles g	473	28	63.6	464	1	08BFR_MOUSE	08BY6_mus_musculu
401	28	63.6	237	1	1BP4_SHEEP	028893_ovis_aries	474	28	63.6	465	2	07XIX6_ORYSA	07XIX6_oryza_sativ
402	28	63.6	239	2	054K45_DICDI	054K45_dicyostelli	475	28	63.6	471	2	05U3U4_BRARE	05U3U4_brachydanio
403	28	63.6	248	2	08X6T7_ECOS7	08X6T7_escherichia	476	28	63.6	484	2	07Y0G5_ORYSA	07Y0G5_oryza_sativ
404	28	63.6	254	1	1BP4_MOUSE	047879_mus_musculu	477	28	63.6	475	2	07ULG9_RHOBA	07ULG9_rhodopirell
405	28	63.6	254	1	1BP4_RAT	P21744_rattus norv	478	28	63.6	491	2	09A7T1_CAUCR	09A7T1_caulobacter
406	28	63.6	254	1	0946U5_HORVU	0946U5_hordewul	479	28	63.6	493	2	051P74_MAGGR	051P74_magnaporthe
407	28	63.6	254	2	08BSM9_MOUSE	08BSM9_mus_musculu	480	28	63.6	507	2	08R0K2_MOUSE	08R0K2_mus_musculu
408	28	63.6	254	2	06B7U4_RAT	06B7U4_rattus norv	481	28	63.6	508	2	07S178_ORYSA	07S178_oryza_sativ
409	28	63.6	258	1	1BP4_BOVIN	005716_bos_taurus	482	28	63.6	509	2	07Q0M5_ANOGA	07Q0M5_anopheles g
410	28	63.6	258	1	1BP4_HUMAN	P22692_homo_sapien	483	28	63.6	521	2	09BAL0_RHIL0	09BAL0_rhizobium l
411	28	63.6	258	2	05U0I2_HUMAN	05U0I2_homo_sapien	484	28	63.6	526	2	04FKN6_TPRYP	04FKN6_rypanosoma
412	28	63.6	268	2	09W1S2_DROME	09W1S2_drosophila	485	28	63.6	531	2	09BZG5_HUMAN	09BZG5_homo_sapien
413	28	63.6	270	2	054201_DICDI	054201_dicyostelli	486	28	63.6	535	2	07R3T0_GJALA	07R3T0_giardia lam
414	28	63.6	272	2	069079_HHVI	069079_human_hep	487	28	63.6	536	2	04NNP1_9DEL1	04NNP1_aeromonxob
415	28	63.6	276	2	04Q092_DROME	04Q092_drosophila	488	28	63.6	538	2	05USC7_HUMAN	05USC7_homo_sapien
416	28	63.6	276	2	08LJP5_GLAMI	08LJP5_antirrhinum	489	28	63.6	539	2	09NUA2_HUMAN	09NUA2_homo_sapien
417	28	63.6	276	2	04TTH5_TERNG	04TTH5_tetradodon n	490	28	63.6	542	2	09BZG6_HUMAN	09BZG6_homo_sapien
418	28	63.6	278	2	06K1W1_ORYSA	06K1W1_oryza_sativ	491	28	63.6	544	2	09BZG7_HUMAN	09BZG7_homo_sapien
419	28	63.6	282	2	082RB2_STRAW	082RB2_streptomyce	492	28	63.6	544	2	05CWN1_ARATH	05CWN1_arabidopsis
420	28	63.6	284	2	054B88_DICDI	054B88_dicyostelli	493	28	63.6	552	2	083RV8_BRADJ	083RV8_bradynhizob
421	28	63.6	288	2	001848_PNECA	001848_pneumocysti	494	28	63.6	557	2	0968Z6_DICDI	0968Z6_dicyostelli
422	28	63.6	293	2	07PMK7_ANOGA	07PMK7_anopheles g	495	28	63.6	559	2	09VZ44_DROME	09VZ44_drosophila
423	28	63.6	294	2	09RCF3_STROCA	09RCF3_streptomyce	496	28	63.6	560	2	054GF7_DICDI	054GF7_dicyostelli
424	28	63.6	296	2	09UUF3_PNECA	09UUF3_pneumocysti	497	28	63.6	562	2	04RHZ7_TERNG	04RHZ7_tetradodon n
425	28	63.6	299	2	04Q1A5_LEIMA	04Q1A5_leishmania	498	28	63.6	567	1	ARS_STRPU	ARS473_strongyloce
426	28	63.6	305	2	09YPF7_HHVI	09YPF7_human_hep	499	28	63.6	569	2	09RFF2_RHOSH	09RFF2_rhodobacter
427	28	63.6	311	2	0580R2_9TTRYP	0580R2_rypanosoma	500	28	63.6	572	2	06NB73_RHOPA	06NB73_rhodospheuo
428	28	63.6	315	2	08KR56_ESCPE	08KR56_escherichia	501	28	63.6	575	2	081RB8_DROME	081RB8_drosophila
429	28	63.6	315	2	0910J4_STROCA	0910J4_streptomyce	502	28	63.6	588	2	0580V3_TPRYP	0580V3_rypanosoma
430	28	63.6	315	2	04S672_TERNG	04S672_tetradodon n	503	28	63.6	591	2	073P17_TREDE	073P17_treponea d
431	28	63.6	316	2	08ZHB3_YERPE	08ZHB3_yersinia pe	504	28	63.6	594	2	084Z97_ORYSA	084Z97_oryza_sativ
432	28	63.6	316	2	0666J2_YERPS	0666J2_yersinia ps	505	28	63.6	594	2	0833X0_ENTFA	0833X0_enterococcu
433	28	63.6	327	2	08D6C0_VIBVU	08D6C0_vibrio vuln	506	28	63.6	599	2	04QB90_LEIMA	04QB90_leishmania
434	28	63.6	328	2	04RPF5_TERNG	04RPF5_tetradodon n	507	28	63.6	603	2	08ZG41_STRAW	08ZG41_streptomyce
435	28	63.6	331	2	057KM1_SALCH	057KM1_salmonella	508	28	63.6	617	2	054NF7_DICDI	054NF7_dicyostelli
436	28	63.6	331	2	05PFO0_SALPA	05PFO0_salmonella	509	28	63.6	619	2	09YMI9_DROME	09YMI9_drosophila
437	28	63.6	331	2	08Z4E2_SALTY	08Z4E2_salmonella	510	28	63.6	624	1	SUYH4_ARATH	088Z6_arabidopsis
438	28	63.6	331	2	08ZML1_SALTY	08ZML1_salmonella	511	28	63.6	628	1	STY1_TEAAT	P3980_saccharomyc
439	28	63.6	334	2	09NZS8_HUMAN	09NZS8_homo_sapien	512	28	63.6	628	2	06RRCO_YEAST	06RRCO_saccharomyc
440	28	63.6	334	2	06DIE4_ERMCT	06DIE4_erynia car	513	28	63.6	628	2	0619B9_CABER	0619B9_caenorhabdi
441	28	63.6	336	2	04J3I4_AZOVI	04J3I4_azocbacter	514	28	63.6	632	1	GBRT_HUMAN	09un88_homo_sapien
442	28	63.6	340	1	LRP1_HHVI	P17888_human_hep	515	28	63.6	632	2	06WGM0_9PARA	06WGM0_mossman vir
443	28	63.6	354	2	061AK0_CABER	061AK0_caenorhabdi	516	28	63.6	634	2	06CX82_KLUTA	06CX82_kluyveromyc
444	28	63.6	355	2	07S6V6_NEUCR	07S6V6_neutrospora	517	28	63.6	648	2	04S516_TERNG	04S516_tetradodon n
445	28	63.6	360	2	09H6B9_HUMAN	09H6B9_homo_sapien	518	28	63.6	654	2	04Q5A1_LEIMA	04Q5A1_leishmania
446	28	63.6	364	2	08ECB2_SHEON	08ECB2_shevanella	519	28	63.6	655	2	06BV47_DBBHA	06BV47_debaryomyce
447	28	63.6	364	2	07YMW5_CABEL	07YMW5_caenorhabdi	520	28	63.6	663	2	09VTX2_DROME	09VTX2_drosophila
448	28	63.6	368	2	04VAA3_MOUSE	04VAA3_mus_musculu	521	28	63.6	668	2	05D7J4_BRABE	05D7J4_brachydanio
449	28	63.6	375	2	094699_PAPRP	094699_parmecium	522	28	63.6	672	2	04IK90_GIBBE1	04IK90_gibberella
450	28	63.6	375	2	04RMC1_TERNG	04RMC1_tetradodon n	523	28	63.6	679	2	08PCTT_XANAC	08PCTT_xanthomonas
451	28	63.6	379	2	08Z3I4_CHLCV	08Z3I4_chlamydomophi	524	28	63.6	696	2	053RS8_HUMAN	053RS8_homo_sapien
452	28	63.6	381	2	09N5Y3_CABEL	09N5Y3_caenorhabdi	525	28	63.6	693	2	09H8R3_HUMAN	09H8R3_homo_sapien
453	28	63.6	383	2	06Z543_GEOCY	06Z543_geodia cydo	526	28	63.6	700	2	06TVG0_9POXY	06TVG0_bovine papu
454	28	63.6	384	2	05Z5R8_MAGGR	05Z5R8_magnaporthe	527	28	63.6	703	2	05S1X6_CRYNE	05S1X6_cryptococcu
455	28	63.6	385	2	075GJ0_ORYSA	075GJ0_oryza_sativ	528	28	63.6	703	2	05KCS7_CRYNE	05KCS7_cryptococcu
456	28	63.6	387	2	06L5H6_ORYSA	06L5H6_oryza_sativ	529	28	63.6	705	2	04N0S5_9DEL1	04N0S5_aeromonxob
457	28	63.6	401	2	06NZ11_BRARE	06NZ11_brachydanio	530	28	63.6	706	2	086H21_DICDI	086H21_dicyostelli
458	28	63.6	411	2	060WF9_CABER	060WF9_caenorhabdi	531	28	63.6	716	2	06TVU2_9POXY	06TVU2_orf_virus.
459	28	63.6	412	2	04J3I5_AZOVI	04J3I5_azocbacter	532	28	63.6	719	2	09UB84_9CIL1	09UB84_parmecium
460	28	63.6	420	2	06FK82_CANGA	06FK82_candida gla	533	28	63.6	720	2	06TW72_9POXY	06TW72_orf_virus.
461	28	63.6	428	2	0786D3_MAGGR	0786D3_magnaporthe	534	28	63.6	721	2	06MM20_BDBBA	06MM20_bdellovibrio
462	28	63.6	428	2	0786D3_9ASPA	0786D3_vanilla aph	535	28	63.6	745	2	081YA6_HUMAN	081YA6_homo_sapien
463	28	63.6	435	2	0616G8_CABER	0616G8_caenorhabdi	536	28	63.6	750	2	054VS4_DICDI	054VS4_dicyostelli
464	28	63.6	437	2	0960H1_HUMAN	0960H1_homo_sapien	537	28	63.6	756	2	054VS2_DICDI	054VS2_dicyostelli
465	28	63.6	437	2	08HDA5_HUMAN	08HDA5_homo_sapien	538	28	63.6	786	2	041IJ2_GIBBE2	041IJ2_gibberella
466	28	63.6	447	2	08N6O3_HUMAN	08N6O3_homo_sapien	539	28	63.6	801	2	05G7N6_9CAUD	05G7N6_listonella
467	28	63.6	445	1	OPCA_STYMP7	054709_synchococc	540	28	63.6	803	1	C13AA_BACTU	054735_bacillus th
468	28	63.6	445	2	05N1G1_SYNP6	05n1g1_synchococc	541	28	63.6	809	2	05S2D5_CRYNE	05S2D5_cryptococcu
469	28	63.6	447	2	054VW4_DICDI	054VW4_dicyostelli	542	28	63.6				

543	28	63.6	809	2	Q5KNQ1_CRYNE	Q5knq1_cryptococcu	616	28	63.6	2319	2	Q0H614_MAIZE	Q0h614_zea mays (m
544	28	63.6	833	2	Q6J288_ACAHA	Q6j288_acanthamoeb	617	28	63.6	2395	2	Q27167_PARTE	Q27167 paramecium
545	28	63.6	873	2	Q9LP01_ARATH	Q9lp01_arabidopsis	618	28	63.6	2397	2	Q6Q7Y4_PAPPR	Q6q7y4_paramecium
546	28	63.6	877	2	Q8WZU1_ORYSA	Q8wzu1_oryza sativ	619	28	63.6	2533	2	P90589_PARTE	P90589 paramecium
547	28	63.6	895	1	ANDR_MACPU	Q97552 macaca faec	620	28	63.6	2533	2	Q27183_PARTE	Q27183 paramecium
548	28	63.6	895	1	ANDR_MACPU	Q6q755 macaca faec	621	28	63.6	2543	2	P90649_PAPPR	P90649 paramecium
549	28	63.6	895	1	ANDR_PAPHA	Q97960 papio hamad	622	28	63.6	2551	2	STAB2_HUMAN	Q8wq48 h stablilin-
550	28	63.6	906	2	Q9UNZ1_HUMAN	Q9unz1_homo sapien	623	28	63.6	2704	1	G168_PAPPR	P17053 paramecium
551	28	63.6	911	1	ANDR_PANTR	Q97775 pan troglod	624	28	63.6	2715	1	G156_PAPPR	P18387 paramecium
552	28	63.6	919	1	ANDR_HUMAN	P10275 homo sapien	625	28	63.6	2717	2	Q94710_PARTE	Q94710 paramecium
553	28	63.6	920	2	Q5UTN9_HUMAN	Q5utn9_homo sapien	626	28	63.6	2729	2	Q6PQK6_PARTE	Q6pqk6_paramecium
554	28	63.6	930	2	Q9NY77_HUMAN	Q9ny77_homo sapien	627	28	63.6	3443	2	Q8JZM6_MOUSE	Q8jzm6_mus musculu
555	28	63.6	930	2	Q7R5C2_GIALA	Q7r5c2_giardia lam	628	27	61.4	34	1	TX1_STRCP	P60931 stromatocopei
556	28	63.6	966	2	Q7UXM4_RHOBA	Q7uxm4_rhodopirell	629	27	61.4	36	1	TUT1B_HADVE	P82227 hadronyche
557	28	63.6	975	2	Q5WRLO_CABEL	Q5wrl0_caenorhabdi	630	27	61.4	36	1	TUT1B_HADVE	P82227 hadronyche
558	28	63.6	980	2	Q24780_CLOTMP	Q24780_clostridium	631	27	61.4	37	1	TUT1C_HADVE	P82228 hadronyche
559	28	63.6	1000	2	Q4FKF7_9TRYP	Q4fkf7_trypanosoma	632	27	61.4	38	2	Q6VRS0_THRCC	P6vrs0_theobroma c
560	28	63.6	1027	2	Q6FWT4_CANGA	Q6fw4_candida gla	633	27	61.4	41	1	TX482_HYSGI	P56854 hyelocrota
561	28	63.6	1035	1	CDC68_YEAST	P32558 saccharomyc	634	27	61.4	43	2	Q6VQF7_CRAVI	Q6vqp7_cirsosotrea
562	28	63.6	1051	2	Q5U4U1_XENLA	Q5u4u1_xenopus lae	635	27	61.4	47	2	GRAM_CRAAB	P01542 cirsambe abys
563	28	63.6	1084	2	Q7SM18_ORYSA	Q7sm18_oryza sativ	636	27	61.4	52	1	Q9N9H0_9BIVA	Q9n9h0_venerupis p
564	28	63.6	1088	1	MMT1_HORVU	Q9mbc2 hordeum vul	637	27	61.4	54	2	Q967H8_9BIVA	Q967h8_cirsosotrea
565	28	63.6	1091	1	MMT1_MAIZE	Q8w519_zea mays (m	638	27	61.4	55	2	Q7E2U4_ORYSA	Q7e2u4_oryza sativ
566	28	63.6	1115	2	Q4UH06_THRAN	Q4uh06_thelateria a	639	27	61.4	55	2	Q7LZAS_9PIPI	Q7lzas_xenopus sp.
567	28	63.6	1125	2	Q9NY75_HUMAN	Q9ny75_homo sapien	640	27	61.4	57	2	Q9N9H2_9BIVA	Q9n9h2_venerupis (
568	28	63.6	1128	2	Q4S6G8_TETNG	Q4s6g8_tetradon n	641	27	61.4	64	2	Q9N9H1_9BIVA	Q9n9h1_venerupis (
569	28	63.6	1139	2	Q54Z39_DICDI	Q54z39_dicyostell	642	27	61.4	64	2	Q9PH34_XYLEFA	Q9ph34_xylella fas
570	28	63.6	1146	2	Q4R8T6_MACPA	Q4r8t6_macaca faec	643	27	61.4	69	2	Q84VGS_ARATH	Q84vgs_arabidopsis
571	28	63.6	1156	2	Q9S938_HUMAN	Q9s938_homo sapien	644	27	61.4	71	2	Q8WQ16_OSTED	Q8wq16_ostrea edul
572	28	63.6	1160	2	Q4RX89_TETNG	Q4rx89_tetradon n	645	27	61.4	71	2	Q80K29_9PAPI	Q80k29_human papil
573	28	63.6	1172	2	Q9LPE05_ARATH	Q9lpe05_arabidopsis	646	27	61.4	72	1	CXAS2_CONST	P28879 conus stria
574	28	63.6	1176	2	Q9NY76_HUMAN	Q9ny76_homo sapien	647	27	61.4	73	2	Q6S9M5_ORYSA	Q6s9ms_oryza sativ
575	28	63.6	1193	2	Q90819_CHICK	Q90819_gallus gall	648	27	61.4	73	2	Q4RA96_TETNG	Q4ra96_tetradon n
576	28	63.6	1202	1	JAG2_RAT	P97607 rattus norv	649	27	61.4	74	1	MT_CRAVI	P23038 cirsosotrea
577	28	63.6	1213	1	JAG1B_BRARE	Q90y54 brachydanio	650	27	61.4	74	2	Q8WQ15_OSTED	Q8wq15_ostrea edul
578	28	63.6	1214	2	Q9MMG2_BLAGB	Q9mmg2_blatteila g	651	27	61.4	74	2	Q6ZFLO_BURMA	Q6zfl0_burholderi
579	28	63.6	1214	2	Q90YD2_XENLA	Q90yd2_xenopus lae	652	27	61.4	75	2	Q6VQP8_CRAVI	Q6vqp8_cirsosotrea
580	28	63.6	1216	2	Q5TEK7_BRARE	Q5tek7_brachydanio	653	27	61.4	75	2	Q9UN15_CRAVI	Q9un15_cirsosotrea
581	28	63.6	1216	2	Q90Y55_BRARE	Q90y55_brachydanio	654	27	61.4	75	2	Q53ZD1_CRAVI	Q53zd1_cirsosotrea
582	28	63.6	1218	1	JAG1_HUMAN	P78504_homo sapien	655	27	61.4	76	2	Q8MUZ7_CRAVI	Q8muz7_cirsosotrea
583	28	63.6	1218	1	JAG1_MOUSE	Q9qxx0_mus musculu	656	27	61.4	76	2	Q6S9S5_ORYSA	Q6s9s5_oryza sativ
584	28	63.6	1218	1	Q4KMR2_HUMAN	Q4kmr2_homo sapien	657	27	61.4	76	2	Q7ULQ4_RHOBA	Q7ulq4_rhodopirell
585	28	63.6	1219	1	JAG1_RAT	Q6j722 rattus norv	658	27	61.4	80	2	Q9BIV4_9BIVA	Q9biv4_cirsosotrea
586	28	63.6	1238	1	JAG1A_BRARE	Q90y57 brachydanio	659	27	61.4	83	2	Q7JVC4_DROME	Q7jvc4_drosophila
587	28	63.6	1242	1	JAG1A_BRARE	Q90y57 brachydanio	660	27	61.4	86	2	Q4GYK6_9TRYP	Q4gyk6_trypanosoma
588	28	63.6	1245	2	Q9Y7V5_TRITHA	Q9y7v5_trichoderma	661	27	61.4	86	2	Q8CLD4_YERPE	Q8cl4_yersinia pe
589	28	63.6	1246	2	Q6LA40_HUMAN	Q6la40_homo sapien	662	27	61.4	86	2	Q6E889_YERPE	Q6e889_yersinia ps
590	28	63.6	1247	1	JAG2_MOUSE	Q9qy65_mus musculu	663	27	61.4	87	2	Q54212_STRDR	Q54212 streptomyce
591	28	63.6	1254	2	Q5TEK8_BRARE	Q5tek8_brachydanio	664	27	61.4	88	2	Q5Z9S6_ORYSA	Q5z9s6_oryza sativ
592	28	63.6	1254	2	Q90Y56_BRARE	Q90y56_brachydanio	665	27	61.4	89	2	Q94376_CABEL	Q94376_caenorhabdi
593	28	63.6	1254	2	Q9YHU2_BRARE	Q9yhu2_brachydanio	666	27	61.4	90	2	Q6L475_ORYSA	Q6l475_oryza sativ
594	28	63.6	1318	2	Q5DIC9_HUMAN	Q5dic9_homo sapien	667	27	61.4	91	2	Q7Y227_9CAUD	Q7y227_siki conver
595	28	63.6	1319	2	Q4S599_CABEL	Q4s599_caenorhabdi	668	27	61.4	92	2	Q8S6B4_9CAUD	Q8s6b4_mycobacteri
596	28	63.6	1374	2	Q6LA41_HUMAN	Q6la41_homo sapien	669	27	61.4	92	2	Q4SB76_TETNG	Q4sb76_tetradon n
597	28	63.6	1403	2	Q9QXG1_MOUSE	Q9qxx1_mus musculu	670	27	61.4	93	2	Q4SCB2_TETNG	Q4scb2_tetradon n
598	28	63.6	1434	2	Q4ITC1_GIBZE	Q4itc1_gibberella	671	27	61.4	94	2	Q6VQPS_CRAVI	Q6vqps_cirsosotrea
599	28	63.6	1466	2	Q5DID0_HUMAN	Q5did0_homo sapien	672	27	61.4	95	2	Q5WQGO_9Z2Z2	Q5wqgo_unclured
600	28	63.6	1469	2	Q6IGM7_CABBR	Q6igm7_caenorhabdi	673	27	61.4	97	2	Q6S6S1_9VIRU	Q6s6s1_lunbo virus
601	28	63.6	1477	2	Q4H3A4_CIOIN	Q4h3a4_ciona intes	674	27	61.4	97	2	Q88472_9VIRU	Q88472_tatyna viru
602	28	63.6	1592	2	Q4IDJ7_GIBZE	Q4idj7_gibberella	675	27	61.4	97	2	Q98751_9VIRU	Q98751_tatyna viru
603	28	63.6	1664	2	Q9TVQ2_CABEL	Q9tvq2_caenorhabdi	676	27	61.4	99	2	Q7D937_MYCTO	Q7d937_mycobacteri
604	28	63.6	1691	2	Q4HUY8_GIBZE	Q4huy8_gibberella	677	27	61.4	99	2	Q7D937_MYCTO	Q7d937_mycobacteri
605	28	63.6	1709	2	Q80XK8_MOUSE	Q80xk8_mus musculu	678	27	61.4	100	2	Q7AJV1_ACRBH	Q7ajv1_agrobacteri
606	28	63.6	1717	2	Q2S566_SCIMA	Q2s566_schistosoma	679	27	61.4	100	2	Q8R4C3_CRICR	Q8r4c3_cricetulus
607	28	63.6	1805	2	Q6S6B1_RAT	Q6s6b1_rattus norv	680	27	61.4	101	2	Q8BQER_MOUSE	Q8bqer_mus musculu
608	28	63.6	1866	2	Q6GQV6_MOUSE	Q6gqv6_mus musculu	681	27	61.4	103	2	Q9AUB8_9ILLI	Q9aub8_copeinicia
609	28	63.6	1896	2	Q84MRS_ORYSA	Q84mrs_oryza sativ	682	27	61.4	104	2	Q9NCR5_9CUCU	Q9ncr5_dendroides
610	28	63.6	1968	2	Q8XOC5_NEUCR	Q8xoc5_neuropora	683	27	61.4	105	2	Q6B715_RABIT	Q6b715_oryctolagus
611	28	63.6	2113	2	Q7XMB7_ORYSA	Q7xmb7_oryza sativ	684	27	61.4	105	2	Q29198_PIG	Q29198_sus scrofa
612	28	63.6	2117	2	Q9H4D8_HUMAN	Q9h4d8_homo sapien	685	27	61.4	107	2	Q9NG19_CRAGI	Q9ng19_cirsosotrea
613	28	63.6	2125	2	Q7QWMI_GIALA	Q7qwm1_giardia lam	686	27	61.4	110	2	Q6IWZ8_9CALI	Q6iwz8_porcine ent
614	28	63.6	2167	2	Q9H4B1_HUMAN	Q9h4b1_homo sapien	687	27	61.4	110	2	Q6IWZ9_9CALI	Q6iwz9_porcine ent
615	28	63.6	2169	2	Q9NY09_HUMAN	Q9ny09_homo sapien	688	27	61.4	112	2	O16120_TENMO	O16120_tenebrio mo

689	27	61.4	112	2	Q9U745_TENNO	Q9U745_tenebrio mo	762	27	61.4	135	2	Q8S4D5_9LILI	Q8s4d5 physokentia
690	27	61.4	112	2	Q9U746_TENNO	Q9u746_tenebrio mo	763	27	61.4	135	2	Q8S4E8_HYOLA	Q8s4e8 hyphorbe 1
691	27	61.4	112	2	Q9U747_TENNO	Q9u747_tenebrio mo	764	27	61.4	135	2	Q9AU72_9LILI	Q9au72 ptychospem
692	27	61.4	112	2	Q9U748_TENNO	Q9u748_tenebrio mo	765	27	61.4	135	2	Q6AB87_PROAC	Q6ab87 propionibac
693	27	61.4	113	2	Q5SNV5_ORYSA	Q5snas_oryza sativ	766	27	61.4	136	2	Q7X6D0_ORYSA	Q7x6d0 oryza sativ
694	27	61.4	115	1	PLAC8_HUMAN	Q9afit1 homo sapien	767	27	61.4	136	2	Q9AU86_9LILI	Q9au86 cyphospema
695	27	61.4	115	2	Q9REJ4_HUMAN	Q96ej4 homo sapien	768	27	61.4	137	2	Q6B941_9PAPI	Q6b941 bovine papl
696	27	61.4	115	2	Q5REK4_PONPY	Q5rek4 pongo pygma	769	27	61.4	137	2	Q8BND5_HUMAN	Q8bnds homo sapien
697	27	61.4	118	2	Q9S979_CRAAB	Q9s979 crambe abys	770	27	61.4	137	2	Q9AU67_9LILI	Q9au67 rectiphiala
698	27	61.4	118	2	Q68531_PSEAC	Q68531 pseudomonas	771	27	61.4	137	2	Q9AU68_9LILI	Q9au68 secretea ex
699	27	61.4	118	2	Q68020_PSEAE	Q68020 pseudomonas	772	27	61.4	137	2	Q9AU69_9LILI	Q9au69 rhopalosytl
700	27	61.4	119	2	Q7UMM3_RHOBA	Q7umms3 rhodopirell	773	27	61.4	137	2	Q9AU70_9LILI	Q9au70 reinhardtia
701	27	61.4	119	2	Q8BAJ5_MOUSE	Q8bj5 mus musculu	774	27	61.4	137	2	Q9AU71_RAVRI	Q9au71 ravenna tiv
702	27	61.4	120	1	VEA_HP42	Q9ytr25 human papil	775	27	61.4	137	2	Q9AU73_PESBC	Q9au73 pseudophoen
703	27	61.4	120	2	Q9Y8R9_AERPE	Q9ytr25 aeropyrum p	776	27	61.4	137	2	Q9AU75_9LILI	Q9au75 ptafeteta e
704	27	61.4	120	2	Q5GCM8_9CUCU	Q5gcm8 microdera d	777	27	61.4	137	2	Q9AU77_9LILI	Q9au77 ilinospadix
705	27	61.4	120	2	Q5G6W9_9CUCU	Q5g6w9 microdera d	778	27	61.4	137	2	Q9AU78_9LILI	Q9au78 ignanura wa
706	27	61.4	120	2	Q8L930_ARATH	Q8l930 arabidopsis	779	27	61.4	137	2	Q9AU82_DYPLE	Q9au82 dypsis lept
707	27	61.4	121	1	Q8L930_ARATH	Q8l930 arabidopsis	780	27	61.4	137	2	Q9AU83_9LILI	Q9au83 dictyosperm
708	27	61.4	121	1	RNPA_NEIGA	Q5f4w3 neisseria g	781	27	61.4	137	2	Q9AU84_9LILI	Q9au84 deckenia no
709	27	61.4	121	1	RNPA_NEIMA	Q9jxk6 neisseria m	782	27	61.4	137	2	Q9AU85_9LILI	Q9au85 cytosactachy
710	27	61.4	121	1	RNPA_NEIMB	Q9jxk6 neisseria m	783	27	61.4	137	2	Q9AU89_CHAHU	Q9au89 chamaetrops
711	27	61.4	121	2	Q7POT7_ANOGA	Q7pct7 anopheles g	784	27	61.4	137	2	Q9AU92_CARMT	Q9au92 cayota mit
712	27	61.4	123	2	Q5X5B3_LEGPA	Q5x5b3 legionella	785	27	61.4	137	2	Q9AU93_9LILI	Q9au93 brongnariari
713	27	61.4	123	2	Q7QF28_ANOGA	Q7qf28 anopheles g	786	27	61.4	137	2	Q9AU94_9LILI	Q9au94 bentinckia
714	27	61.4	124	2	Q45261_CAEBL	Q45261 caenorhabdl	787	27	61.4	137	2	Q9AU95_9LILI	Q9au95 beccatiopho
715	27	61.4	124	2	Q16121_TENNO	Q16121 tenebrio mo	788	27	61.4	137	2	Q69GY1_9PAPI	Q69gy1 bovine papl
716	27	61.4	125	2	Q6DX5_TENNO	Q6dx5 tenebrio mo	789	27	61.4	138	2	Q9AU81_9LILI	Q9au81 dypsis lute
717	27	61.4	125	2	Q6C9J7_MOUSE	Q6c9j7 mus musculu	790	27	61.4	143	2	Q5KOP2_ORYSA	Q5kpf2 oryza sativ
718	27	61.4	127	2	Q9AU74_9LILI	Q9au74 prestoea ca	791	27	61.4	143	2	Q6QNU6_9PAPI	Q6qnu6 human papil
719	27	61.4	127	2	Q9AU76_9LILI	Q9au76 phoenicopho	792	27	61.4	144	2	Q8ALU7_922Z2	Q8alj7 uncultured
720	27	61.4	128	2	Q9YDS2_AERPE	Q9yds2 aeropyrum p	793	27	61.4	144	2	Q6S4F5_9PAPI	Q6s4f5 human papil
721	27	61.4	128	2	Q8BNU7_MOUSE	Q8bnu7 mus musculu	794	27	61.4	144	2	Q6S4I3_9PAPI	Q6s4i3 human papil
722	27	61.4	129	2	Q9AU90_CHACS	Q9au90 chamaedorea	795	27	61.4	144	2	Q80K25_9PAPI	Q80k25 human papil
723	27	61.4	130	2	Q8S4D3_9LILI	Q8s4d3 rhopaloblas	796	27	61.4	144	2	Q8J5K9_9PAPI	Q8j5k9 chimpanzee
724	27	61.4	130	2	Q8S4D9_9LILI	Q8s4d9 normandya n	797	27	61.4	144	2	Q8QNU7_9PAPI	Q8qnu7 human papil
725	27	61.4	130	2	Q9T040_ARATH	Q9t040 arabidopsis	798	27	61.4	144	2	Q8QR28_9PAPI	Q8qr28 human papil
726	27	61.4	130	2	Q9BGV7_RHILLO	Q9bgv7 rhizobium l	799	27	61.4	144	2	Q8QR29_9PAPI	Q8qr29 human papil
727	27	61.4	131	2	Q5T7M1_HUMAN	Q5t7m1 homo sapien	800	27	61.4	144	2	Q8VLR6_9PAPI	Q8vlr6 human papil
728	27	61.4	131	2	Q7XJY2_ORYSA	Q7xjy2 oryza sativ	801	27	61.4	144	2	Q9JGX9_9PAPI	Q9jgx9 human papil
729	27	61.4	131	2	Q9AU79_9LILI	Q9au79 heterospath	802	27	61.4	145	1	PA21B_LATSE	Pa21b laticauda s
730	27	61.4	132	2	Q8S4F5_9LILI	Q8s4f5 actinorhytl	803	27	61.4	145	2	Q6VOP2_CRAVI	Q6vop2 crassostrea
731	27	61.4	132	2	Q9AU97_9LILI	Q9au97 bacillus gas	804	27	61.4	145	2	Q7D250_AGRIS	Q7d250 agrobacteri
732	27	61.4	133	2	Q6ZBR3_ORYSA	Q6zbr3 oryza sativ	805	27	61.4	145	2	Q8UJZ4_9PAPI	Q8juz4 macaca faec
733	27	61.4	133	2	Q9AU80_9LILI	Q9au80 sclerospem	806	27	61.4	145	2	Q8V919_9PAPI	Q8v919 human papil
734	27	61.4	133	2	Q9AU80_9LILI	Q9au80 tetradon n	807	27	61.4	145	2	Q6ZOW3_HUMAN	Q6zow3 homo sapien
735	27	61.4	134	2	Q4SDH4_TENNO	Q4sdh4 tetradon n	808	27	61.4	146	2	Q80K31_9PAPI	Q80k31 human papil
736	27	61.4	134	2	Q5Q0B1_ARATH	Q5q0b1 arabidopsis	809	27	61.4	146	2	Q8V9J1_9PAPI	Q8v9j1 human papil
737	27	61.4	134	2	Q8S4C6_9LILI	Q8s4c6 wellia regl	810	27	61.4	146	2	Q9WGF3_9PAPI	Q9wgf3 human papil
738	27	61.4	134	2	Q8S4C8_9LILI	Q8s4c8 sommeria e	811	27	61.4	147	2	Q8UJL6_9PAPI	Q8jul6 chimpanzee
739	27	61.4	134	2	Q8S4C9_9LILI	Q8s4c9 sclerospem	812	27	61.4	147	2	Q8UJZ3_9PAPI	Q8juz3 macaca faec
740	27	61.4	134	2	Q8S4D0_9LILI	Q8s4d0 satkentia	813	27	61.4	147	2	Q9DMY6_9PAPI	Q9dmy6 human papil
741	27	61.4	134	2	Q8S4D1_ROTRE	Q8s4d1 roysetonea r	814	27	61.4	148	2	Q9NCQ8_9CUCU	Q9ncq8 dendroides
742	27	61.4	134	2	Q8S4D2_9LILI	Q8s4d2 roschertia m	815	27	61.4	148	2	Q16122_TENNO	Q16122 tenebrio mo
743	27	61.4	134	2	Q8S4D4_9LILI	Q8s4d4 podococcus	816	27	61.4	148	2	Q6H5T8_ORYSA	Q6h5t8 oryza sativ
744	27	61.4	134	2	Q8S4D6_9LILI	Q8s4d6 pelagodoxa	817	27	61.4	148	2	Q9ERK8_MOUSE	Q9erk8 mus musculu
745	27	61.4	134	2	Q8S4D7_9LILI	Q8s4d7 orania tris	818	27	61.4	149	2	Q6VOP3_CRAVI	Q6vop3 crassostrea
746	27	61.4	134	2	Q8S4D8_9LILI	Q8s4d8 oncosperma	819	27	61.4	149	2	Q6VOP4_CRAVI	Q6vop4 crassostrea
747	27	61.4	134	2	Q8S4E1_9LILI	Q8s4e1 nenga pumil	820	27	61.4	150	2	Q7RU43_NEUCR	Q7ru43 neurospora
748	27	61.4	134	2	Q8S4E2_9LILI	Q8s4e2 masojala mad	821	27	61.4	151	1	NGB_PIG	Ngb pig
749	27	61.4	134	2	Q8S4E3_9LILI	Q8s4e3 macrojaya d	822	27	61.4	152	2	Q4PIW5_CAEBL	Q4piw5 caenorhabdl
750	27	61.4	134	2	Q8S4E5_9LILI	Q8s4e5 lodolicea ma	823	27	61.4	154	2	Q7R3E7_GIALA	Q7r3e7 gibberia lam
751	27	61.4	134	2	Q8S4E6_9LILI	Q8s4e6 leopoldinia	824	27	61.4	155	2	Q4J3J8_GIBZBE	Q4j3j8 gibberella
752	27	61.4	134	2	Q8S4E7_9LILI	Q8s4e7 lemmrophoen	825	27	61.4	155	2	Q6ZS77_HUMAN	Q6zs77 homo sapien
753	27	61.4	134	2	Q8S4E9_9LILI	Q8s4e9 gytrophyllu	826	27	61.4	155	2	Q5DCT3_SCHJA	Q5dct3 schistosoma
754	27	61.4	134	2	Q8S4F0_9LILI	Q8s4f0 gausilia may	827	27	61.4	155	2	Q4KCK4_PSEFS	Q4kck4 pseudonons
755	27	61.4	134	2	Q8S4F2_9LILI	Q8s4f2 cyrtosachy	828	27	61.4	155	2	Q5PFY8_SALPA	Q5pfy8 salmonella
756	27	61.4	134	2	Q8S4F1_9LILI	Q8s4f1 dypsis hete	829	27	61.4	155	2	Q8Z916_SALTI	Q8z916 saltonella
757	27	61.4	134	2	Q8S4F4_9LILI	Q8s4f4 areca catec	830	27	61.4	157	2	Q4TEP1_TENNG	Q4tef1 tetradon n
758	27	61.4	134	2	Q9AU96_BAUSE	Q9au96 balaka seem	831	27	61.4	161	2	Q6NTG6_HUMAN	Q6ntg6 homo sapien
759	27	61.4	134	2	Q8UJ11_AGRIS	Q8uj11 agrobacteri	832	27	61.4	163	2	Q6K6S2_ORYSA	Q6k6s2 oryza sativ
760	27	61.4	135	1	H32_XENIA	P02302 xenopus lae	833	27	61.4	164	2	Q4H1M0_9BACT	Q4h1m0 uncultured
761	27	61.4	135	1	H32_XENIA	P02302 xenopus lae	834	27	61.4	164	2	Q4H1M0_9BACT	Q4h1m0 uncultured

835	27	61.4	167	2	Q5D5H7_9RICK	Q5d5h7 wolbachia e	908	27	61.4	244	1	RS6_BRAFL	O01727 branchiosto
836	27	61.4	167	2	Q8UYL3_9POXV	Q8uy13 vaccinia vl	909	27	61.4	245	2	O14644_HUMAN	O14644 homo sapien
837	27	61.4	167	2	Q9DJL2_COMPX	Q9dj12 compox viru	910	27	61.4	245	2	Q75LR5_ORYSA	Q75LR5 oryza sativ
838	27	61.4	167	2	Q72762_COMPX	Q72762 compox viru	911	27	61.4	246	1	Y208_METUA	Y208 METUA
839	27	61.4	168	2	Q60BZ3_ORYSA	Q60Bz3 oryza sativ	912	27	61.4	247	1	RS6_APLCA	Q9bmx5 aplysia cal
840	27	61.4	168	2	Q5RQ72_GLUOX	Q5rQ72 gluconbact	913	27	61.4	247	2	Q6SPR3_SOTBN	Q6bpr3 glycine max
841	27	61.4	169	2	Q72768_HUMAN	Q72768 homo sapien	914	27	61.4	247	2	Q4VBAB_RAT	Q4vbab rattus norv
842	27	61.4	169	2	Q9JKE0_RAT	Q9jke0 rattus norv	915	27	61.4	247	2	Q9DID3_MOUSE	Q9did3 mus musculu
843	27	61.4	170	2	Q6LGV4_PHOPR	Q6lGV4 photobacter	916	27	61.4	247	2	Q91VTI_MOUSE	Q91vti mus musculu
844	27	61.4	172	1	Y3408_IACPPL	Q88sk6 lactobacill	917	27	61.4	247	2	Q9PTD6_CHICK	Q9ptd6 gallus gall
845	27	61.4	173	1	NUCM_FELSU	Q796f9 pelomedusa	918	27	61.4	248	1	RS6_DROME	P29327 drosophila
846	27	61.4	175	2	Q9FM63_ARATH	Q9fm63 arabidopsis	919	27	61.4	248	2	Q4RG40_TETNG	Q4rg40 tetradon n
847	27	61.4	175	2	Q53P21_ORYSA	Q53p21 oryza sativ	920	27	61.4	249	1	RS6B_ARATH	P51430 arabidopsis
848	27	61.4	176	2	Q4J7H8_SULAC	Q4j7h8 sulfolobus	921	27	61.4	249	1	RS6_CHICK	P47838 gallus gall
849	27	61.4	177	2	Q6J112_DROME	Q6j112 drosophila	922	27	61.4	249	1	RS6_HUMAN	P62753 homo sapien
850	27	61.4	177	2	Q9MSY0_ORYSA	Q9msy0 oryza sativ	923	27	61.4	249	1	RS6_ICTPU	Q90y78 ictalurus p
851	27	61.4	177	2	Q9MSY7_ORYSA	Q9msy7 oryza sativ	924	27	61.4	249	1	RS6_MOUSE	P62754 mus musculu
852	27	61.4	179	2	Q7Y061_ORYSA	Q7y061 oryza sativ	925	27	61.4	249	1	RS6_ONCMY	Q9Y912 oncorhynch
853	27	61.4	183	1	RM43_MOUSE	Q99h89 mus musculu	926	27	61.4	249	1	RS6_RAT	P62755 rattus norv
854	27	61.4	186	2	Q6AMR6_DROME	Q6amr6 drosophila	927	27	61.4	249	1	RS6_XENLA	P39017 xenopus lae
855	27	61.4	186	2	Q9VAL2_DROME	Q9val2 drosophila	928	27	61.4	249	2	Q8N6Z7_HUMAN	Q8n6z7 homo sapien
856	27	61.4	188	2	Q9L031_STRCO	Q9l031 streptomyce	929	27	61.4	249	2	Q9EDV6_HUMAN	Q9edv6 homo sapien
857	27	61.4	188	2	Q9ERS8_RAT	Q9er88 rattus norv	930	27	61.4	249	2	Q4VBV7_HUMAN	Q4vby7 homo sapien
858	27	61.4	190	2	Q65637_ARATH	Q65637 arabidopsis	931	27	61.4	249	2	Q5CXZ9_CRYPO	Q5cxz9 cryptospori
859	27	61.4	190	2	P70368_MOUSE	P70368 mus musculu	932	27	61.4	249	2	Q5CP03_CRYHO	Q5cp03 cryptospori
860	27	61.4	191	2	Q4QFX9_LEIMA	Q4qfx9 leishmania	933	27	61.4	249	2	Q5E995_BOVIN	Q5e995 boe taurus
861	27	61.4	192	2	Q54V14_DICDI	Q54v14 dictyostell	934	27	61.4	249	2	Q4R4K6_MACFA	Q4r4k6 macaca fasc
862	27	61.4	193	2	Q7UM46_RHOBA	Q7um46 rhodospirell	935	27	61.4	249	2	Q7Y0H9_BRANA	Q7y0h9 brassica na
863	27	61.4	195	2	Q55BC9_DICDI	Q55bc9 dictyostell	936	27	61.4	249	2	Q9JLM3_NOSFU	Q9jlm3 nosce punc
864	27	61.4	196	2	Q8VC17_MOUSE	Q8vc17 mus musculu	937	27	61.4	249	2	Q8BT09_MOUSE	Q8bt09 mus musculu
865	27	61.4	197	2	Q7M097_VIBRY	Q7m097 vibrio vuln	938	27	61.4	249	2	Q5BLK1_MOUSE	Q5blk1 mus musculu
866	27	61.4	199	2	Q61XF7_CABER	Q61xf7 caenorhabdi	939	27	61.4	249	2	Q7ZYU0_XENLA	Q7zyu0 xenopus lae
867	27	61.4	200	2	Q6VOP0_CRAVI	Q6vop0 crassostrea	940	27	61.4	249	2	Q5D086_XENLA	Q5d086 xenopus lae
868	27	61.4	201	2	Q8GTL3_TOBAC	Q8gtl3 nicotiana t	941	27	61.4	249	2	Q6DHL6_BRAHE	Q6dhl6 brachydano
869	27	61.4	202	2	Q5EHV1_GECUA	Q5ehv1 gecko japon	942	27	61.4	249	2	Q4SDU7_TETNG	Q4sdj7 tetradon n
870	27	61.4	204	2	Q6VQPI_CRAVI	Q6vqpi crassostrea	943	27	61.4	249	2	Q6P7K2_XENTR	Q6p7k2 xenopus tro
871	27	61.4	209	2	Q7WNM2_BORBR	Q7wnm2 bordetella	944	27	61.4	250	1	RS6A_ARATH	Q46484 corbact
872	27	61.4	210	2	Q6IGL5_DROME	Q6igl5 drosophila	945	27	61.4	250	2	Q8LH97_ORYSA	Q8lh97 oryza sativ
873	27	61.4	210	2	Q4Q3D3_LEIMA	Q4q3d3 leishmania	946	27	61.4	250	2	Q4TH17_TETNG	Q4th17 tetradon n
874	27	61.4	210	2	Q6A8A1_LEIIX	Q6a8a1 leifsonia x	947	27	61.4	251	1	RS6_ASPOP	Q9m3v8 asparagila o
875	27	61.4	211	2	Q8BR56_MOUSE	Q8br56 mus musculu	948	27	61.4	251	2	Q8IRN9_DROME	Q8irn9 drosophila
876	27	61.4	211	2	Q8R037_MOUSE	Q8r037 mus musculu	949	27	61.4	251	2	Q04014_MAIJE	Q04014 zea mays (m
877	27	61.4	213	2	Q5F3Y3_CHICK	Q5f3y3 gallus gall	950	27	61.4	251	2	Q9EY50_MAIJE	Q9ey50 zea mays (m
878	27	61.4	214	2	Q7SSU1_NEUCR	Q7ssu1 neurospora	951	27	61.4	252	2	Q5OX16_ENTHI	Q5ox16 entamoeba h
879	27	61.4	215	2	Q8VKP8_MYCTU	Q8vkp8 mycobacteri	952	27	61.4	253	1	ERMA_CORDI	P16898 corynebacte
880	27	61.4	219	2	Q9P3P8_NEUCR	Q9p3f8 neurospora	953	27	61.4	253	1	RS6_MANSE	Q94624 manduca sex
881	27	61.4	220	1	SODP_CAMUR	P53f40 campylobact	954	27	61.4	253	1	RS6_SPOFR	Q9sv32 epdoptera
882	27	61.4	220	2	Q5HMZ7_CAMUR	Q5hmz7 campylobact	955	27	61.4	253	2	Q5UAN8_BOWMO	Q5uan8 bombyx mori
883	27	61.4	222	2	Q6A854_PROAC	Q6a854 propionibac	956	27	61.4	253	2	Q46484_CORBT	Q46484 corynebacte
884	27	61.4	223	2	Q4WBL0_ASPPU	Q4wbl0 aspergillus	957	27	61.4	253	2	Q53WT1_CORDI	Q53wt1 corynebacte
885	27	61.4	223	2	Q86YK5_HUMAN	Q86y55 homo sapien	958	27	61.4	253	2	Q7BBX6_PROAC	Q7bbx6 propionibac
886	27	61.4	226	2	Q8A4U3_HUMAN	Q8a4u3 bacteroides	959	27	61.4	254	2	Q6TAY3_CHICK	Q6tay3 gallus gall
887	27	61.4	229	2	Q9NKM5_HUMAN	Q9nm5 homo sapien	960	27	61.4	255	1	TNR9_HUMAN	Q07011 homo sapien
888	27	61.4	229	2	Q57289_HUMAN	Q57289 homo sapien	961	27	61.4	255	2	Q9H5H2_HUMAN	Q9h5h2 homo sapien
889	27	61.4	231	1	NRG1_YEAST	Q03125 saccharomye	962	27	61.4	255	2	Q94LS9_ORYSA	Q94ls9 oryza sativ
890	27	61.4	232	2	Q9SKA7_MACFA	Q9skf9 macaca fasc	963	27	61.4	256	1	TNR9_MOUSE	P20334 mus musculu
891	27	61.4	233	2	Q8BFW6_XANAC	Q8bfw9 xanthomonas	964	27	61.4	256	2	Q7RRG5_PLAFO	Q7rrc5 plasmodium
892	27	61.4	234	2	Q7Z6G5_HUMAN	Q7z6g5 homo sapien	965	27	61.4	256	2	Q4YUS6_PLAAB	Q4yus6 plasmodium
893	27	61.4	234	2	Q6YSD4_ORYSA	Q6yrd4 oryza sativ	966	27	61.4	256	2	Q4XZM0_PLACH	Q4xzm0 plasmodium
894	27	61.4	234	2	Q9X897_STRCO	Q9x897 streptomyce	967	27	61.4	256	2	Q8F0Z4_LEPIN	Q8f0z4 leptospira
895	27	61.4	235	2	Q90ZF9_CHICK	Q90zf9 gallus gall	968	27	61.4	257	2	Q9AM12_STRAT	Q9am12 streptomyce
896	27	61.4	236	1	SRTD1_HUMAN	Q9uhv2 homo sapien	969	27	61.4	258	2	Q72U49_LEPIC	Q72u49 leptospira
897	27	61.4	236	1	SRTD1_MOUSE	Q9j110 mus musculu	970	27	61.4	258	2	Q4W8J3_RAT	Q4w8j3 rattus norv
898	27	61.4	236	2	Q53GCO_HUMAN	Q53gco homo sapien	971	27	61.4	258	2	Q4V895_RAT	Q4v895 rattus norv
899	27	61.4	236	2	Q8XWK3_RALSO	Q8xwk3 talstonia s	972	27	61.4	259	2	Q93519_CABEL	Q93519 caenorhabdi
900	27	61.4	236	2	Q6P771_RAT	Q6p771 rattus norv	973	27	61.4	259	2	Q63PV9_BURPS	Q63pv9 burkholderi
901	27	61.4	237	1	IPDE_DICDI	P22549 dictyostell	974	27	61.4	259	2	Q7NKM5_GLOVI	Q7ncm5 gloeobacter
902	27	61.4	237	2	Q54YU3_DICDI	Q54y43 dictyostell	975	27	61.4	260	2	Q8LN62_ORYSA	Q8ln62 oryza sativ
903	27	61.4	237	1	Q45804_CABEL	Q45804 caenorhabdi	976	27	61.4	263	2	Q6CXB9_KLUHA	Q6cxb9 kluyveromyc
904	27	61.4	239	2	GAG_AVEVI	P06936 avian endog	977	27	61.4	265	2	Q9A2T7_CAUCR	Q9a2t7 culicobacter
905	27	61.4	239	2	Q4KTD6_SUBDO	Q4ktd6 superlites d	978	27	61.4	265	2	Q7Z6V5_DBSVA	Q7z6v5 deauleovibr
906	27	61.4	239	2	Q4ZF55_CLOPE	Q4zf55 clostridium	979	27	61.4	267	2	Q02764_RABIT	Q02764 oryctolagus
907	27	61.4	242	2	Q6JWW6_GUROC	Q6jww6 cistopileura	980	27	61.4	270	2	Q75SV8_FELCA	Q75sv8 felis silve

```
981 27 61.4 270 2 Q9ZUN5 ARATH
982 27 61.4 270 2 Q9LH57 ARATH
983 27 61.4 270 2 Q8BX32 MOUSE
984 27 61.4 271 1 TNR4_RAT
985 27 61.4 271 2 Q54RT0_DICDI
986 27 61.4 271 2 Q5PQP5_RAT
987 27 61.4 272 1 TNR4_MOUSE
988 27 61.4 272 1 Q6Z5Z9_ORYSA
989 27 61.4 273 2 Q5UCF4_MAI2E
990 27 61.4 274 1 TNR5_CANPA
991 27 61.4 274 1 Q4IVQ6_AZOTI
992 27 61.4 274 2 Q8P3B6_XANCP
993 27 61.4 275 2 Q8BQ83_HUMAN
994 27 61.4 275 2 Q8QW09_MOUSE
995 27 61.4 276 2 Q4IFD8_GIB2E
996 27 61.4 276 2 Q94473_DICDI
997 27 61.4 276 2 Q71F55_MOUSE
998 27 61.4 277 1 TNR5_HUMAN
999 27 61.4 277 1 TNR5_HUMAN
1000 27 61.4 277 2 Q5U007_HUMAN
```

ALIGNMENTS

RESULT 1

Q98AR9_RH10 PRELIMINARY; PRT; 192 AA.

```
ID Q98AR9_RH10 PRELIMINARY; PRT; 192 AA.
AC Q98AR9;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE M15880 protein.
GN OrderedLocustNames=mlr5880;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
NCBI_TaxID=381;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti."
RL EMBL: BA000012; BAB52253.1; -; Genomic_DNA.
DR HSSP: Q92FY9; IPTH.
DR GO: GO:0016491; P:oxidoreductase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR002347; Adh_short_C2.
DR InterPro: IPR002424; Insect_adh_fam.
DR PRINTS: PR00081; GDHRD.
DR PRINTS: PR01167; INSADHFAMILY.
DR PRINTS: PR00080; SDRFAMILY.
DR Complete proteome.
KW SEQUENCE 192 AA; 20183 MW; FD2F660D156037BC CRC64;
```

Query Match 81.8%; Score 36; DB 2; Length 192;
Best Local Similarity 71.4%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 176 CSPATGC 182

RESULT 2

```
Q86AK7_DICDI
ID Q86AK7_DICDI PRELIMINARY; PRT; 360 AA.
AC Q86AK7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Similar to Dictyostelium discoidium (Slime mold). Prescalk protein
DE (Hypothetical protein).
ORFNames=DDB0168042; DDB0217347;
GN Dictyostelium discoidium (Slime mold).
OS Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
NCBI_TaxID=44689;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szefranski K., Pachebat J.A.,
RA Bankier A.T., Dear P.H., Lehmann R., Baumgart C., Parra G.,
RA Abiri J.F., Guigo R., Kump K., Tunggal B., Cox E., Quail M.A.,
RA Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoidium."
RL Nature 418:79-85(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RC Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RC Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugang R., Barriam M., Song J., Olsen R., Szefranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerthout P., Desany B., Just E., Worio T., Kost R., Churcher C.,
RA Cooper J., Haydock S., Van Driessche N., Cronin A., Goodhead I.,
RA Muzy D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Qules M., Mohan M.B., Saito T., Buchliesser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Loulsegue H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Splegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winkler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA Chiscolm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Batelli B., Kuspa A.;
RT "The genome of the social amoeba Dictyostelium discoidium."
RL Nature 0:0-0(2005).
RL EMBL: AC116984; AA051380.1; -; Genomic_DNA.
DR EMBL: AAF101000025; EAL70624.1; -; Genomic_DNA.
DR EMBL: AAF101000023; EAL70779.1; -; Genomic_DNA.
DR HSSP: P01382; INTN.
DR InterPro: IPR002172; LDL_receptor_A.
DR InterPro: IPR001673; S_mold_repeat.
DR Pfam: PF00526; Dicty_CTCG; 13.
DR PRINTS: PR00261; LDLRECEPTOR.
KW Hypothetical protein.
KW SEQUENCE 360 AA; 38013 MW; 4B0061ABA26ED81E CRC64;
```

Query Match 81.8%; Score 36; DB 2; Length 360;
Best Local Similarity 71.4%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 211 CSPSTGC 217

RESULT 3

Q8KJ72_RH10 PRELIMINARY; PRT; 546 AA.

```

AC 08XJ72;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE PUTATIVE SHORT-CHAIN TYPE DEHYDROGENASE/REDUCTASE PROTEIN.
GN Name=ms1329;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=R7A;
RX MEDLINE=2199272; PubMed=12003951;
RX DOI=10.1126/JB.184.11.3086-3095.2002;
RA Sullivan J.T., Trebicki J.R., Crickshank R.W., Gouzy J.,
RA Brown S.D., Elliot R.M., Fleetwood D.J., McCallum N.G., Rosbach U.,
RA Stuart G.S., Weaver J.E., Webb R.J., de Bruijn F.J., Ronsom C.W.;
RT "Comparative sequence analysis of the symbiosis island of
RT Mesorhizobium loti strain R7A."
RL J. Bacteriol. 184:3086-3095(2002).
CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SPR) family.
DR EMBL: AL672114; CAD31361.1; -; Genomic_DNA.
DR HSSP; P50163; ZAE1.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; Adh_short.
DR InterPro; IPR002347; Adh_short_C2.
DR InterPro; IPR002424; Insect_adh_fam.
DR Pfam; PF00106; adh_short; 2.
DR PRINTS; PR00081; GDHRDH.
DR PRINTS; PR01167; INSADHFAMILY.
DR PRINTS; PR00080; SDRPAMILY.
DR PROSITE; PS00061; ADH_SHORT; 1.
KW Oxidoreductase.
SQ
SEQUENCE 546 AA; 56900 MW; 091D2EFDE855A9C CRC64;

Query Match
Best Local Similarity 81.8%; Score 36; DB 2; Length 546;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 530 CSPATGC 536

RESULT 4
054X44 D1CD1 PRELIMINARY; PRT; 597 AA.
ID 054X44 D1CD1 PRELIMINARY;
AC 054X44;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DD80205657;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugang R., Barriman M., Song J., Olsen R., Szatranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivoero F.,
RA Bankler A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kertomou A., Nie X., Hall N., Anjard C., Hemphill L., Bacon N.,
RA Fairbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangaveilu M., Johnson D., Knights A.,

```

```

RA Louisedge H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shalaby G., Schlieper M., Weinstein G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzner M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Bärrell B., Kuspa A.;
RT "The genome of the social amoeba Dictyostelium discoideum."
RL Nature 0:0-0(2005).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAF10100059; EAL67758.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ
SEQUENCE 597 AA; 61409 MW; AA97F1D0234D336C CRC64;

Query Match
Best Local Similarity 81.8%; Score 36; DB 2; Length 597;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 500 CSPPTGC 506

RESULT 5
07MAJ3 D1CD1 PRELIMINARY; PRT; 601 AA.
ID 07MAJ3 D1CD1 PRELIMINARY;
AC 07MAJ3;
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE DIF-induced prestalk pbd63 protein precursor (Fragments).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=87187613; PubMed=3568124; DOI=10.1016/0092-8674(87)90559-9;
RA Williams J.G., Ceccarelli A., McRobbie S., Mahubani H., Kay R.R.,
RA Early A., Berks M., Jermyn K.A.;
RT "Direct induction of Dictyostelium prestalk gene expression by DIF
RT provides evidence that DIF is a morphogen."
RL Cell 49:185-192(1987).
DR PIR; A27020; A27020.
DR InterPro; IPR001673; S_mold_repeat.
DR Pfam; PF00526; Dicty_CTDC; 19.
FT NON_TER 601
SQ
SEQUENCE 601 AA; 63359 MW; 7D4433616CDAC438 CRC64;

Query Match
Best Local Similarity 81.8%; Score 36; DB 2; Length 601;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 245 CSPSTGC 251

RESULT 6
0627A0 CAEBR PRELIMINARY; PRT; 608 AA.
ID 0627A0 CAEBR PRELIMINARY;
AC 0627A0;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG00767 (Fragment).
GN Name=CBG00767;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]

```

RP NUCLEOTIDE SEQUENCE.
RG The C bridge Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC01000006; CAES7750.1; -; Genomic_DNA.
CC CC
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006150; Worn_repeat_1.
DR SMART; SM00289; WRI; 6.
DR PROSITE; PS01186; EGF_2; UNKNOWN_2.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 608 AA; 65898 MW; F09R2470D/BE6393 CRC64;

Query Match 81.8%; Score 36; DB 2; Length 608;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10
Db 205 CAPSTGC 211

RESULT 7
OSAYG2_DICDI PRELIMINARY; PRT; 1710 AA.
ID OSAYG2;
AC OSAYG2;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Extracellular matrix protein STR430.
GN Name=ecma; ORFNames=DOB0220137;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=XX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugang R., Berriman M., Song J., Olsen R., Szatranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bacon N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., Van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Louisgeed H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Splegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzner M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Bartelli B., Kuspa A.;
RT "The genome of the social amoeba Dictyostelium discoideum";
RT Nature 0:0-0 (2005).
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAF101000052; EAL6809.1; -; Genomic_DNA.
KW Matrix protein.
SQ SEQUENCE 1710 AA; 177828 MW; 008C47B85EFDD7D2 CRC64;

Query Match 81.8%; Score 36; DB 2; Length 1710;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10
Db 245 CSPSTGC 251

RESULT 8
ID OS5K02_MACFA PRELIMINARY; PRT; 148 AA.
AC OS5K02;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecoinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Medulla oblongata;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062948; BAB60737.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 148 AA; 15473 MW; C8D2AA301E0C8191 CRC64;

Query Match 79.5%; Score 35; DB 2; Length 148;
Best Local Similarity 71.4%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10
Db 74 CGPSTGC 80

RESULT 9
ID OS9N018_MACFA PRELIMINARY; PRT; 148 AA.
AC OS9N018;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecoinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cerebellum cortex;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
libraries";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046629; BAB03547.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 148 AA; 15495 MW; 6242BCB430C6E13B CRC64;

Query Match 79.5%; Score 35; DB 2; Length 148;
Best Local Similarity 71.4%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10
Db 74 CGPSTGC 80

RESULT 10
ID OS5BT6_DICDI PRELIMINARY; PRT; 370 AA.
AC OS5BT6;
DT 13-SEP-2005 (TREMBLrel. 31, Created)

DT 13-SEP-2005 (TReMBLrel. 31, last sequence update)
 DT 13-SEP-2005 (TReMBLrel. 31, last annotation update)
 DE Hypothetical protein.
 GN ORFNames=DD0191005;
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 ON NCB1_TaxID=44689;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AX4;
 RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
 RA Sugeng R., Berriman M., Song J., Olsen R., Szatranski K., Xu Q.,
 RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
 RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
 RA Kerhroux A., Nie X., Hall N., Anjard C., Hemphill L., Bacon N.,
 RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
 RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,
 RA Hauser H., James K., Quiles M., Mohan M.B., Salto T., Buchrieser C.,
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
 RA Louised H., Mungall K., Oliver K., Price C., Quail M.A.,
 RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
 RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
 RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,
 RA "The genome of the social amoeba Dictyostelium discoideum.";
 RL Nature 0:0-0(2005).
 CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL:AF10100011; EAL72534.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 370 AA; 40291 MW; FB3DF46B49A751A CRC64;

QY 4 CXPXTGC 10
 Db 293 CNPSTGC 299

Query Match 79.5%; Score 35; DB 2; Length 370;
 Best Local Similarity 71.4%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 11
 Q55BX0 D1CD1 PRELIMINARY; PRT; 370 AA.
 AC Q55BX0;
 DT 13-SEP-2005 (TReMBLrel. 31, Created)
 DT 13-SEP-2005 (TReMBLrel. 31, last sequence update)
 DT 13-SEP-2005 (TReMBLrel. 31, last annotation update)
 DE Hypothetical protein.
 GN ORFNames=DD0190973;
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 ON NCB1_TaxID=44689;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AX4;
 RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
 RA Sugeng R., Berriman M., Song J., Olsen R., Szatranski K., Xu Q.,
 RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
 RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
 RA Kerhroux A., Nie X., Hall N., Anjard C., Hemphill L., Bacon N.,
 RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
 RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,
 RA Hauser H., James K., Quiles M., Mohan M.B., Salto T., Buchrieser C.,
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
 RA Louised H., Mungall K., Oliver K., Price C., Quail M.A.,

RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
 RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
 RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,
 RA "The genome of the social amoeba Dictyostelium discoideum.";
 RL Nature 0:0-0(2005).
 CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL:AF10100011; EAL72514.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 370 AA; 40278 MW; 1E904E2F1186E25 CRC64;

QY 4 CXPXTGC 10
 Db 293 CNPSTGC 299

Query Match 79.5%; Score 35; DB 2; Length 370;
 Best Local Similarity 71.4%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 12
 Q55KP3 CRYNE PRELIMINARY; PRT; 466 AA.
 AC Q55KP3;
 DT 10-MAY-2005 (TReMBLrel. 30, Created)
 DT 10-MAY-2005 (TReMBLrel. 30, last sequence update)
 DT 10-MAY-2005 (TReMBLrel. 30, last annotation update)
 DE Expressed protein.
 GN ORFNames=CNC02330;
 OS Cryptococcus neoformans var. neoformans JEC21.
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
 OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
 ON NCB1_TaxID=214684;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=JEC21;
 RA Loftus B., Amedeo P., Roncaglia P., Vamathavan J., Uterback T.,
 RA Van Aken S., Fraser C.,
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=JEC21;
 RX PubMed=15653466; DOI=10.1126/science.1103773;
 RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
 RA Vamathavan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
 RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
 RA D'Souza C.A., Fox D.S., Grubberg V., Fu J., Fukushima M., Haas B.J.,
 RA Huang J.C., Janbon G., Jones S.J.M., Kdo H.L., Kryzinski M.I.,
 RA Kwon-Chung K.J., Lengele K.B., Mailli R., Maira W.A., Maira R.B.,
 RA Mathewson C.A., Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L.,
 RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
 RA Suh B.B., Tenney A., Uterback T.R., Wicks B.L., Wortman J.R.,
 RA Wye N.H., Kronstad J.W., Lodge J.K., Heltman J., Davis R.W.,
 RA Fraser C.M., Hyman R.W.,
 RT "The genome of the basidiomycetous yeast and human pathogen
 RT Cryptococcus neoformans.";
 RL Science 307:1321-1324(2005).
 DR EMBL:AB017343; AAM42205.1; -; Genomic_DNA.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00018; SH3.1; 1.
 DR PRINTS: PR00452; SH3DOMAIN.
 KW Complete proteome.
 SQ SEQUENCE 466 AA; 50328 MW; 8D287B46E77EE27 CRC64;

QY 4 CXPXTGC 10

Query Match 79.5%; Score 35; DB 2; Length 466;
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 408 CDPSTGC 414

RESULT 13
Q9A0P5_PSERE PRELIMINARY; PRT; 571 AA.

AC Q9A0P5_PSERE PRELIMINARY; PRT; 571 AA.
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative ABC transporter subunit.
GN Name=ORF31;

OS Pseudomonas resinovorans.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=53412;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CA10;
RX MEDLINE=21264379; PubMed=11371531;
DOI=10.1128/JB.183.12.3663-3679.2001;

RA Nojiri H., Sekiguchi H., Maeda K., Urata M., Nakai S., Yoshida T.,
Habe H., Omori T.;
RT "Genetic characterization and evolutionary implications of car gene
cluster in carbazole-degrader, Pseudomonas sp. strain CA10."
RJ J. Bacteriol. 183:3663-3679(2001).
RN [2]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CA10;
RX MEDLINE=97386424; PubMed=9244273;
Sato S., Onchiyama N., Kimura T., Nojiri H., Yamane H., Omori T.;

RT "Cloning of genes involved in carbazole degradation of Pseudomonas sp.
strain CA10: nucleotide sequence of genes and characterization of
meta-cleavage enzymes and hydrolase."
RJ J. Bacteriol. 179:4841-4849(1997).
RN [3]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CA10;
RX MEDLINE=97386425; PubMed=9244274;
Sato S., Nam J., Kasuga K., Nojiri H., Yamane H., Omori T.;

RT "Identification and characterization of genes encoding carbazole 1,9a-
dioxygenase in Pseudomonas sp. strain CA10."
RJ J. Bacteriol. 179:4850-4858(1997).
RN [4]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CA10;
RX PubMed=15466034; DOI=10.1128/JB.186.20.6815-6823.2004;
Urata M., Miyakoshi M., Kai S., Maeda K., Habe H., Omori T.,

Yamane H., Nojiri H.;
RT "transcriptional regulation of the ant operon, encoding two-component
antitranilactate 1,2-dioxygenase, on the carbazole-degradative plasmid
PCAR1 of Pseudomonas resinovorans strain CA10."
RJ J. Bacteriol. 186:6815-6823(2004).
RN [5]

DR EMBL; AB047548; BAB32742.1; -; Genomic_DNA.
GO; GO:0016020; C:membrane; IEA.
GO; GO:0005215; F:transporter activity; IEA.
GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001851; Bac_inmem_transp.
DR Pfam; PF02653; BPD_transp_2; 2.

DR EMBL; AB047548; BAB32742.1; -; Genomic_DNA.
GO; GO:0016020; C:membrane; IEA.
GO; GO:0005215; F:transporter activity; IEA.
GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001851; Bac_inmem_transp.
DR Pfam; PF02653; BPD_transp_2; 2.

DR EMBL; AB047548; BAB32742.1; -; Genomic_DNA.
GO; GO:0016020; C:membrane; IEA.
GO; GO:0005215; F:transporter activity; IEA.
GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001851; Bac_inmem_transp.
DR Pfam; PF02653; BPD_transp_2; 2.

ID Q4SBN7_TETNG PRELIMINARY; PRT; 997 AA.

AC Q4SBN7_TETNG PRELIMINARY; PRT; 997 AA.
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 15 SCAF1467, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00020889001;

OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
Maucci E., Bouneau L., Fischer C., Ozouf-Costat C., Bernot A.,
Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
Rathouard V., Jubin C., Castelli V., Katinka M., Vachier B.,
Biemont C., Skallil Z., Cattolico L., Poulain J., De Bernardis V.,
Cruaud C., Duprat S., Brothier P., Coutanceau J.P., Gouzy J.,
Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
Landet V., Schachler V., Queller F., Saurin W., Scarpelli C.,
Mincker P., Lander E.S., Weissbach J., Roost Crolius H.;

RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype."
RJ Nature 431:946-957(2004).
RN [2]

RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
Submitted (Feb-2004) to the EMBL/GenBank/DBJ databases.
RL -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01014667; CAG01945.1; -; Genomic_DNA.

FT NON_TER 1 997
FT NON_TER 1 997
SQ SEQUENCE 997 AA; 114462 MW; 60278BFD562DEBC CRC64;

Query Match 79.5%; Score 35; DB 2; Length 997;
Best Local Similarity 71.4%; Pred. No. 8.6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 293 CSPITGC 299

RESULT 15
Q4SFS2_TETNG PRELIMINARY; PRT; 1018 AA.

AC Q4SFS2_TETNG PRELIMINARY; PRT; 1018 AA.
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAF14608, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00019252001;

OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
Maucci E., Bouneau L., Fischer C., Ozouf-Costat C., Bernot A.,
Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,

RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype."
RJ Nature 431:946-957(2004).
RN [2]

RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
Maucci E., Bouneau L., Fischer C., Ozouf-Costat C., Bernot A.,
Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,

RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castellino L., Katinka M., Vacherie B.,
 RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Bernardis V.,
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
 RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
 RA Kellis M., Volt J.N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Landt V., Schacher V., Quetier F., Saurin W., Scarpelli C.,
 RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.,
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 the early vertebrate proto-karyotype.";
 RL Nature 431:946-957(2004).
 [2]
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope; Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC EMBL; CA001014608; CAG00730.1; -; Genomic_DNA.
 CC InterPro; IPR006209; EGF like.
 DR InterPro; IPR002049; Laminin EGF.
 DR InterPro; IPR008211; Laminin_N.
 DR InterPro; IPR001134; Netrin_C.
 DR Pfam; PF00053; Laminin_EGF; 3.
 DR Pfam; PF01759; NTR; 1.
 DR SMART; SM00643; C345C; 1.
 DR SMART; SM00180; EGF_Lam; 3.
 DR SMART; SM00136; LAMNT; 1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE; PS01248; EGF_LAM_1; 3.
 DR PROSITE; PS50027; EGF_LAM_2; 3.
 DR PROSITE; PS5117; LAMININ_NTR; 1.
 DR PROSITE; PS50189; NTR; 1.
 KM Laminin EGF-like domain.
 FT NON_TER
 FT NON_TER 1018 1018
 FT SEQUENCE 1018 AA; 113875 MW; A47C021431A85FF4 CRC64;
 SQ

Query Match 79.5%; Score 35; DB 2; Length 1018;
 Best Local Similarity 71.4%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
 DB 770 CSDITGC 776

RESULT 16
 ID 045795_TETNG PRELIMINARY; PRT; 1839 AA.
 AC 045795;
 DT 13-SEP-2005 (TREMBlrel. 31, Created)
 DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
 DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
 DE Chromosome 1 SCAF14716, whole genome shotgun sequence.
 DE (Fragment).
 GN ORFNames=GSTENG00022912001;
 OS Tetraodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthopterygii; Acanthopterygii; Percomorphi; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 OC NCB1_Taxid=99883;
 RX [1]_Taxid=99883;
 RP NUCLEOTIDE SEQUENCE.
 RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
 RA Mauceli E., Bouneau L., Fischer C., Orouf-Coetz C., Bernot A.,
 RA Nicard S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Seguens B.,
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castellino L., Katinka M., Vacherie B.,
 RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Bernardis V.,
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,

RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
 RA Kellis M., Volt J.N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Landt V., Schacher V., Quetier F., Saurin W., Scarpelli C.,
 RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.,
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 the early vertebrate proto-karyotype.";
 RL Nature 431:946-957(2004).
 [2]
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope; Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC EMBL; CA001014716; CAG03487.1; -; Genomic_DNA.
 CC InterPro; IPR001952; ALK_phosphatase.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR003887; LEW.
 DR Pfam; PF00245; ALK_phosphatase; 2.
 DR Pfam; PF03020; LEW; 1.
 DR PRINTS; PR00113; ALKPHPTASE.
 DR SMART; SM00098; ALKPC; 2.
 DR SMART; SM00540; LEW; 1.
 DR PROSITE; PS00123; ALKALINE_PHOSPHATASE; UNKNOWN_1.
 DR PROSITE; PS50954; LEW; 1.
 KM Magnesium; Zinc.
 FT NON_TER 1839 1839
 FT SEQUENCE 1839 AA; 203924 MW; 05A14C4AAEBC7BC9 CRC64;
 SQ

Query Match 79.5%; Score 35; DB 2; Length 1839;
 Best Local Similarity 71.4%; Pred. No. 4.6e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
 DB 1418 CTPGTGC 1424

RESULT 17
 ID 072607_HUMAN PRELIMINARY; PRT; 113 AA.
 AC 072607;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE BIP2B5 protein.
 DE BIP2B5 protein.
 GN Name=BIP2B5; GN
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
 OC Homo.
 OC NCB1_Taxid=9606;
 RX [1]_Taxid=9606;
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=lung;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straube R.L., Pelngold R.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Burow K.H., Scheefel C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Dlatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stanleone M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carinci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kertman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,

```

RA Schnerch A., Schein J.E., Jones S.J.M., Maria M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RG NIH MGC Project;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC053646; AAHS3646.1; -; mRNA.
SQ SEQUENCE 113 AA; 11613 MW; 8624A42297FAA3FF CRC64;

Query Match
Best Local Similarity 71.4%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 103 CYPKTC 109

RESULT 18
PA2X NOTSC STANDARD; PRT; 145 AA.
AC P20146;
ID PA2X NOTSC
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Probable phospholipase A2 precursor (BC 3.1.1.4) (Phosphatidylcholine
2-acylhydrolase).
OS Nectechis scutatus scutatus (Mainland tiger snake) (Common tiger
snake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodermata; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Acanthophinae; Nectechis.
OX NCBI_TaxId=70142;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=venom gland;
RA Duceanal F.;
RL Submitted (JAN-1989) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PA2 catalyzes the calcium-dependent hydrolysis of the 2-
acyl groups in 3-sn-phosphoglycerides.
CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O = 1-
acylglycerophosphocholine + a carboxylate.
CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the phospholipase A2 family. Group I
subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
DR EMBL; X14043; CAA32201.1; -; mRNA.
DR PIR; S07983; S07983.
DR HSSP; P00608; 1AE7.
DR InterPro; IPR001211; PhospholipaseA2.
DR PANTHER; PTHR11716; PhospholipaseA2; 1.
DR Pfam; PF00068; Phospholip_A2.1; 1.
DR PRINTS; PR00389; PPHPLIPSEA2.
DR ProDom; PD000303; PhospholipaseA2; 1.
DR SMART; SM00085; PA2c; 1.
DR PROSITE; PS00119; PA2_ASP; 1.
DR PROSITE; PS00118; PA2_HIS; 1.
KW Calcium: Hydrolase; Lipid degradation; Metal-binding;
KW Multigene family; Signal.
FT SIGNAL 1 21 Potential.
FT PROPEP 22 27 Potential.
FT CHAIN 28 145 Probable phospholipase A2.
FT ACT_SITE 75 75 By similarity.

```

```

FT ACT_SITE 119 119 By similarity.
FT METAL 55 55 Calcium (via carbonyl oxygen) (By
FT METAL 57 57 Calcium (via carbonyl oxygen) (By
FT METAL 57 57 similarity).
FT METAL 76 76 Calcium (By similarity).
FT DISULFID 38 98 By similarity.
FT DISULFID 54 144 By similarity.
FT DISULFID 56 72 By similarity.
FT DISULFID 71 125 By similarity.
FT DISULFID 78 118 By similarity.
FT DISULFID 87 111 By similarity.
FT DISULFID 105 116 By similarity.
SQ SEQUENCE 145 AA; 16002 MW; 38E36029AE5FAC9 CRC64;

Query Match
Best Local Similarity 71.4%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 105 CDPKTC 111

RESULT 19
Q7U7V9 SYNXP PRELIMINARY; PRT; 167 AA.
ID Q7U7V9 SYNXP
AC Q7U7V9;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical precursor.
DE Hypothetical precursor.
DE OrderedCusNames=SYNM0871;
OS Synechococcus sp. (Strain WH8102).
OC Synechococcus; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxId=84588;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L.,
RA Chaitin P., Lamerdin J.E., Regala W., Allen E.E., McCarren J.,
RA Paulsen I.T., Dutresne A., Patensky F., Webb E.A., Waterbury J.;
RL "The genome of a motile marine Synechococcus.";
DR EMBL; BX569691; CAE07386.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein; Signal.
FT SIGNAL 1 22 Potential.
FT SEQUENCE 167 AA; 18419 MW; 0B3AFB830ECF971C CRC64;

Query Match
Best Local Similarity 77.3%; Score 34; DB 2; Length 167;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 133 CYPKTC 139

RESULT 20
Q54YF0 D1CDI PRELIMINARY; PRT; 296 AA.
ID Q54YF0 D1CDI
AC Q54YF0;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Hypothetical protein.
DE ORFNames=DDB0218080;
GN Dictyostelium discoideum (Slime mold).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxId=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;

```

RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
 RA Sucgang R., Berriman M., Song J., Olsen R., Szatranski K., Xu Q.,
 RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
 RA Banker A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
 RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
 RA Fairbrother P., Deany B., Just E., Morio T., Rost R., Churcher C.,
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
 RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
 RA Louisedge H., Mungall K., Oliver K., Price C., Quail M.A., Sanders M.,
 RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
 RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
 RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,
 RA "The genome of the social amoeba Dictyostelium discoideum.";
 RT Nature 0:0-0(2005).
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL; AAF10100052; EAL68495.1; -; Genomic_DNA.
 DR Hypothetical protein.
 KW SEQUENCE 296 AA; 33010 MW; A741D9A5252B5479 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 296;
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
 DB 26 CDPNTGC 32

RESULT 21
 O9SUVA ARATH PRELIMINARY; PRT; 312 AA.
 AC O9SUVA3-
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Hypothetical protein F884.80 (Hypothetical protein AT4g32380).
 GN Name=F884.80; Synonyms=AT4g32380;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Bevan M., Terry N., Ardiles W., Buyschaert C., Daseville R.,
 RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,
 RA Villalobos R., Gielens J., Van Montagu M., Hohensei J., Mewes H.W.,
 RA Mayer K.F.X., Lemke K., Schueller C.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA EU Arabidopsis sequencing project;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RA Terry N., Ardiles W., Buyschaert C., Daseville R., De Clerck R.,
 RA De Keyser A., Neyt P., Rouze P., Van Den Daele H., Villalobos R.,
 RA Gielens J., Van Montagu M., Mewes H.W., Lemke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL015457; CAA2565.1; -; Genomic_DNA.
 DR EMBL; AL161581; CAB79955.1; -; Genomic_DNA.

DR PIR; T05348; T05348.
 DR GO; GO:0004650; P:polygalacturonase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR000743; Glyco_hydro_28.
 DR InterPro; IPR006626; PDI.
 DR Pfam; PF00295; Glyco_hydro_28; 2.
 DR SMART; SM00710; PDI; 3.
 KW Hypothetical protein.
 SQ SEQUENCE 312 AA; 34095 MW; E2E70A2622F30BE0 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 312;
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
 DB 208 CDPNTGC 214

RESULT 22
 O54N14 DICDI PRELIMINARY; PRT; 380 AA.
 AC O54N14;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=DD8018573;
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AX4.

RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
 RA Sucgang R., Berriman M., Song J., Olsen R., Szatranski K., Xu Q.,
 RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
 RA Banker A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
 RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
 RA Fairbrother P., Deany B., Just E., Morio T., Rost R., Churcher C.,
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
 RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
 RA Louisedge H., Mungall K., Oliver K., Price C., Quail M.A., Sanders M.,
 RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
 RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
 RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,
 RA "The genome of the social amoeba Dictyostelium discoideum.";
 RT Nature 0:0-0(2005).
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL; AAF101000124; EAL64620.1; -; Genomic_DNA.
 DR Hypothetical protein.
 KW SEQUENCE 380 AA; 40195 MW; A40E4A8D346C8B8D CRC64;

Query Match 77.3%; Score 34; DB 2; Length 380;
 Best Local Similarity 71.4%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
 DB 124 CDPNTGC 130

RESULT 23
 O6TVH9_9POXV PRELIMINARY; PRT; 465 AA.
 ID O6TVH9_9POXV

```

AC 06TVH9;
DT 05-VUL-2004 (Tremblrel. 27, Created)
DT 05-VUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-VUL-2004 (Tremblrel. 27, Last annotation update)
DE Hypothetical protein.
DE Bovine papular stomatitis virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Parapoxvirus.
OC NCBI_TaxID=129727;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BV-AR02;
RX PubMed=14671098; DOI=10.1128/JVI.78.1.168-177.2004;
RA Delton G., Tulman E.R., Alfonso C.L., Lu Z., Piccone M.E., Kutish G.F.,
RA de la Concha-Bernierillo A., Leimkühn H.D., Piccone M.E., Kutish G.F.,
RA Rock D.L.;
RT "Genomes of the Parapoxviruses Orf Virus and Bovine Papular
RT Stomatitis Virus."
RL J. Virol. 78:168-177(2004).
DR EMBL: AY386265; AAR98366.1; -; Genomic_DNA.
DR InterPro: IPR007027; Pox_F11.
DR Pfam: PF04943; Pox_F11; 1.
KW Hypothetical protein.
SQ SEQUENCE 465 AA; 51404 MW; BD735469CBED1879 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 465;
Best Local Similarity 71.4%; Pred. No. 2e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 5; Conservative 0;

OY 4 CXPXTGC 10
| | | |
| | | |
Db 26 CVPRTGC 32

RESULT 24
O4MMCI ASPFU
ID O4MMCI ASPFU PRELIMINARY; PRT; 487 AA.
AC O4MMCI_
DT 13-SEP-2005 (Tremblrel. 31, Created)
DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
DE Chitin deacetylase, putative.
GN ORFNames=Afu6910430;
OS Aspergillus fumigatus Af393.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic trichocomaceae; Aspergillus.
OC NCBI_TaxID=330879;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF293.
RA Nierman W., Pain A., Anderson M.J., Wortman J., Kim H., Stanley,
RA Airova J., Berriman M., Abe K., Archer D.B., Bernerjo C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA Foster N., Fraser A., Garcia J.L., Garcia M.J., Goble A.,
RA Goldman G.H., Goni K., Griffith-Jones S., Gwilliam R., Haas B.,
RA Haas H., Harris D., Horluchi H., Huang J., Humphrey S., Jimenez J.,
RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulikarni R.,
RA Kumagai T., Latton A., Latge J.-P., Li W., Lord A., Lu C.,
RA Majors W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,
RA Mounya I., Mulligan S., Murphy L., O'Neill S., Paulsen I.,
RA Penativa M.A., Perera M., Price C., Pritchard B.L., Quail M.A.,
RA Rabinowitz E., Rawlins N., Rajandream M.-A., Reichard U.,
RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA Rensing C.M., Rutter S., Salzberg S.L., Sanchez M.,
RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA Takeuchi M., Tekle A.F., Turner G., Vazquez de Aldana C.R., Weidman J.,
RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,
RA Machida M., Hall N., Barrett B., Denning D.W.;
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus
RT Aspergillus fumigatus."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an

```

```

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAHP01000006; EAL88893.1; -; Genomic_DNA.
DR InterPro: IPR001002; Chitin_bd_1.
DR InterPro: IPR002509; Polysac_deacet.
DR Pfam: PF00187; Chitin_bind_1; 1.
DR Pfam: PF01522; Polysac_deac_1; 1.
DR Pfam: PD000609; Chitin_binding_1; 2.
DR SMART: SM00270; Chitin_bd_1; 2.
DR PROSITE: PS00941; CHIT_BIND_I_2; 2.
KW Chitin-binding.
SQ SEQUENCE 487 AA; 52742 MW; 1386016F0E98F50 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 487;
Best Local Similarity 71.4%; Pred. No. 2.1e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 5; Conservative 0;

OY 4 CXPXTGC 10
| | | |
| | | |
Db 241 CVPRTGC 247

RESULT 25
O5ZAZ8 ORYSA
ID O5ZAZ8 ORYSA PRELIMINARY; PRT; 504 AA.
AC O5ZAZ8;
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Putative aspartic proteinase nepenthesin II.
GN Name=P0413G02.28;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OC NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22337376; PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nishimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikono M., Ito S., Ito T., Ito Y., Ito Y., Iwabuchi A., Kamiya K.,
RA Kanaeawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Kanchita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takasaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yokawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.-I., Eun M.-Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1."
RL Nature 420:312-316(2002).
DR EMBL: AP003344; BAD53242.1; -; Genomic_DNA.
DR GO: GO:0004194; F:pepsin A activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR001461; Peptidase_A1.
DR Pfam: PF00026; Asp; 1.
DR PRINTS: PR00792; PEPsin.
SQ SEQUENCE 504 AA; 53090 MW; 186FBA9CC5D745D CRC64;

Query Match 77.3%; Score 34; DB 2; Length 504;
Best Local Similarity 71.4%; Pred. No. 2.1e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 5; Conservative 0;

OY 4 CXPXTGC 10
| | | |
| | | |
Db 119 CVPRTGC 125

RESULT 26

```

OS4N64.DICDI
 ID OS4N64.DICDI PRELIMINARY; PRT; 507 AA.
 AC OS4N64;
 DT 13-SEP-2005 (T-EMBLrel. 31, Created)
 DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=DD80218721;
 OS Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
 OC Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
 OX NCBI_TaxId=44689;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AX4;
 RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
 RA Sugeng R., Berrieman M., Song J., Olsen R., Szafanski K., Xu Q.,
 RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
 RA Banker A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
 RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Baeson N.,
 RA Farbrother P., Desany B., Just B., Morio T., Roest R., Churcher C.,
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
 RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
 RA Wastrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
 RA Loulès H., Mangall K., Oliver K., Price C., Quail M.A.,
 RA Uruñehara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
 RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzler M., Kay R.R.,
 RA Williams J., Dear P.H., Noegel A.A., Barrell B., Karp A.;
 RA "The genome of the social amoeba Dictyostelium discoideum.";
 RT Nature 0:0-0(2005).
 RL -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAF10100124; EAL64710.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 507 AA; 53214 MW; F5AFAC31BACEF5 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 507;
 Best Local Similarity 71.4%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
 DB 184 CNPLTGC 190

RESULT 27
 ID OS4SFD9.T-ETNG PRELIMINARY; PRT; 670 AA.
 AC OS4SFD9;
 DT 13-SEP-2005 (T-EMBLrel. 31, Created)
 DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
 DE Chromosome 2 SCAF14604, whole genome shotgun sequence.
 GN ORFNames=GSTENG0019142001;
 OS Tetraodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 OX NCBI_TaxId=99883;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
 RA Mauceli E., Bonneau L., Fischer C., Orouf-Costaz C., Bernot A.,
 RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Dubin C., Castellani V., Katinka M., Vacherie B.,

RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
 RA Cruaud C., Duprat S., Broctier P., Coutancan J.P., Gonzy J.,
 RA Parra G., Lardier G., Chaplet C., McKernan K.J., McEwan P., Bosak S.,
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Laudet V., Schachter V., Queller F., Saulin W., Scarpelli C.,
 RA Winkler P., Lander E.S., Weissbach J., Roest Croollis H.;
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 the early vertebrate proto-karyotype.";
 RL Nature 431:946-957(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope; Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; CA601014604; CAG00643.1; -; Genomic_DNA.
 FT NON TER 1
 SQ SEQUENCE 670 AA; 76484 MW; 0715D5B56262C30 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 670;
 Best Local Similarity 71.4%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
 DB 114 CXPXTGC 120

RESULT 28
 ID OSB3W6.EMENI PRELIMINARY; PRT; 870 AA.
 AC OSB3W6;
 DT 10-MAY-2005 (T-EMBLrel. 30, Created)
 DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (T-EMBLrel. 30, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=AN4764.2;
 OS Aspergillus nidulans FGSC A4.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; EmERICELLA.
 OX NCBI_TaxId=227321;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=FGSC A4;
 RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
 RA Arachchi H.M., Barina N., Bastien V., Bloom T., Boguslavsky L.,
 RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
 RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., Dearellano K.,
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
 RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
 RA Gairyna S., Gierre S., Graham L., Grand-Pierre N., Halez N.,
 RA Hagopian D., Hagoos B., Hall J., Horton L., Hulne W., Iliev I.,
 RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
 RA Kellis C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
 RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
 RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Menais L.,
 RA Mihova T., Milenga V., Murphy T., Naylor T., Nguyen C., Nicol R.,
 RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
 RA Oliver J., Peterson K., Phunhkhang P., Pletre N., Purcell S.,
 RA Rancupha A., Ramsamy U., Raymond C., Retta R., Rise C., Rogov P.,
 RA Roman J., Schuer S., Schupbach R., Sedman S., Severi P., Smirnov S.,
 RA Smith C., Spencer B., Strange-Thomann N., Stojanovic N., Stubbs M.,
 RA Talama J., Teste S., Theodore V., Toppan K., Travers M.,
 RA Vassiliev H., Venkataratan V.S., Viel R., Vo A., Wang S., Wilson B.,
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
 RA Lander E.;
 RT "Genome Sequence of Aspergillus nidulans.";
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

DR EMBL: AACD01000080: EAA60806.1; -: Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 870 AA; 96093 MW; 2679DA01B40B338 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 870;
 Best Local Similarity 71.4%; Pred. No. 3.8e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
 DB 478 CXPXTGC 484

RESULT 29

OS4N02_DICDI PRELIMINARY; PRT; 932 AA.
 AC 054N02;
 DT 13-SEP-2005 (TReMBLrel. 31, Created)
 DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=DDB0186586;
 OS Dictyostelium discoideum (slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AX4;
 RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
 RA Sugang R., Bertrman M., Song J., Olsen R., Szafianski K., Xu Q.,
 RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
 RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
 RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Baeson N.,
 RA Farbrother P., Deasny B., Just E., Morio T., Rost R., Churcher C.,
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
 RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
 RA Louisege H., Mungall K., Oliver K., Price C., Quail M.A.,
 RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
 RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzner M., Kay R.R.,
 RA Williams J., Dear P.H., Noegel A.A., Bartell B., Kupa A.,
 RA "The genome of the social amoeba Dictyostelium discoideum.";
 RL Nature 0:0-0(2005).
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL: AAF101000124; EAL64630.1; -: Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 932 AA; 99556 MW; EF21C56D1CCA5E22 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 932;
 Best Local Similarity 71.4%; Pred. No. 3.8e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
 DB 609 CXPXTGC 615

RESULT 30

OS69K4_DICDI PRELIMINARY; PRT; 998 AA.
 AC 0869K4;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 10-MAY-2005 (TReMBLrel. 30, Last annotation update)
 DE Similar to Mus musculus (Mouse). Tenascin X (Hypothetical
 protein).

GN ORFNames=DDB0167241;
 OS Dictyostelium discoideum (slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AX4;
 RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
 RA Gloeckner G., Eichinger L., Szafianski K., Pachebat J.A.,
 RA Bankier A.T., Dear P.H., Lehmann R., Baumgart C., Parra G.,
 RA Abril J.F., Guigo R., Kump K., Tunggal B., Cox E., Quail M.A.,
 RA Platzner M., Rosenthal A., Noegel A.A.;
 RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
 RL Nature 418:79-85(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AX4;
 RA Baumgart C.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AX4;
 RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
 RA Sugang R., Bertrman M., Song J., Olsen R., Szafianski K., Xu Q.,
 RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
 RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
 RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Baeson N.,
 RA Farbrother P., Deasny B., Just E., Morio T., Rost R., Churcher C.,
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
 RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
 RA Louisege H., Mungall K., Oliver K., Price C., Quail M.A.,
 RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
 RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzner M., Kay R.R.,
 RA Williams J., Dear P.H., Noegel A.A., Bartell B., Kupa A.,
 RA "The genome of the social amoeba Dictyostelium discoideum.";
 RL Nature 0:0-0(2005).
 CC EMBL: AAF101000030; EAL69580.1; -: Genomic_DNA.
 DR EMBL: AAF101000030; EAL69580.1; -: Genomic_DNA.
 DR HSSP: P01132; IEGF.
 DR GO: GO:0008305; C:integrin complex; IEA.
 DR GO: GO:0005198; F:structural molecule activity; IEA.
 DR GO: GO:0007160; P:cell-matrix adhesion; IEA.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR006209; EGF-like.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR002043; Integrin alpha.
 DR InterPro: IPR002049; Laminin_EGF.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF01839; FG-GAP; 1.
 DR PRINTS: PRO0011; EGF_1.
 DR SMART: SM00181; EGF_9.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_10.
 DR PROSITE: PS01186; EGF_2; 7.
 DR PROSITE: PS50026; EGF_3; 8.
 KW Hypothetical protein.
 SQ SEQUENCE 998 AA; 106001 MW; F79BEF394D3E2369 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 998;
 Best Local Similarity 71.4%; Pred. No. 4e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
 DB 609 CXPXTGC 615

RESULT 31

OS5C92_DICDI
ID OS5C92_DICDI PRELIMINARY; PRT; 999 AA.
AC OS5C92;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DE Hypothetical protein.
GN ORFNames=DD80216748;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugeng R., Berriman M., Song J., Olsen R., Szatanski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Kontorov B.A., Rivero F.,
RA Bankler A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bacon N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Louised H., Mungall K., Oliver K., Price C., Quail M.A., Sanders M.,
RA Urushihara H., Hernandez J., Rabinowitz E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.A.,
RA Williams J., Dear P.H., Noegel A.A., Bartlett B., Kuspa A.;
RT "The genome of the social amoeba Dictyostelium discoideum.";
RL Nature 0:0-0(2005).
CC -I- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL:AAFI01000011; EMBL72871.1; -; Genomic_DNA.
CC DR EMBL; AAFI01000011;
CC KW Hypothetical protein.
SQ SEQUENCE 999 AA; 106586 MW; FC387975B85D9C CRC64;

Query Match 77.3%; Score 34; DB 2; Length 999;
Best Local Similarity 71.4%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 609 CNPLTGC 615

RESULT 32
Q57XW5_9TRY PRELIMINARY; PRT; 1147 AA.
ID Q57XW5_9TRY
AC Q57XW5;
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
DE Hypothetical protein.
GN ORFNames=TD927.3.1320;
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
NCBI_TaxID=5691;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Ghedin E., Blandin G., Bartholomew D., Caler E., Haas B., Hannick L.,
RA Shallom J., Hou L., Djikeng A., Feldlyum T., Hostetler J.,
RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
RA Khalak H.G., Salzberg S., Simpson A.J., Tallon L., Van Aken S.,
RA Wanless D., White O., Wortman J., Frazer C.M., El-Sayed N.M.A.;
RL Submitted (Apr-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC159411; AAX69554.1; -; Genomic_DNA.
KW Hypothetical protein.

SQ SEQUENCE 1147 AA; 124779 MW; 60C07617752B4AA7 CRC64;
Query Match 77.3%; Score 34; DB 2; Length 1147;
Best Local Similarity 71.4%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 343 CVPRTGC 349

RESULT 33
Q6Y625_9VIRU PRELIMINARY; PRT; 2217 AA.
ID Q6Y625_9VIRU
AC Q6Y625;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DE Polymerase.
OS Lassa virus.
OC Viruses; ssRNA negative-strand viruses; Arenaviridae; Arenavirus;
OC Old world arenaviruses.
NCBI_TaxID=11620;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CSF;
RA Viet H.S., Gunther S.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY179174; AA059515.1; -; Genomic_RNA.
DR GO; GO:0003968; R:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0016740; R:transferase activity; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR010453; Arena_RNA_pol.
DR InterPro; IPR001000; Glyco_Hydro_10.
DR InterPro; IPR007099; RNA_pol_NSVir.
DR Pfam; PF06317; Arena_RNA_pol; 1.
DR PROSITE; PS00591; GLYCOSYL_HYDROL_F10; UNKNOWN 1.
SQ SEQUENCE 2217 AA; 253347 MW; 577DBF9A5A546EC7 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 2217;
Best Local Similarity 71.4%; Pred. No. 8.5e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 1001 CGPWTGC 1007

RESULT 34
Q4IR98_GIBZE PRELIMINARY; PRT; 77 AA.
ID Q4IR98_GIBZE
AC Q4IR98;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DE Hypothetical protein.
GN ORFNames=RG00260.1;
OS Gibberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
NCBI_TaxID=229533;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arabach H.M., Barina N., Bastien V., Bloom T., Boguslavsky L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collamore A., Cook A., Cooke P., Corum B., Dearellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson U., Fero S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
RA Gardyna S., Gierre S., Graham U., Grand-Pierre N., Hatz N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,

RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
RA Mathews C., Mauceli E., McCarthy M., Meidrum J., Menes L.,
RA Mihova T., Mienga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Oliver A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Rachtupke A., Schauer S., Schuppach R., Seaman S., Severy P., Smirnov S.,
RA Roman J., Schauer S., Schuppach R., Stojanovic N., Stubbs M.,
RA Smith C., Spencer B., Strange-Thomann N., Topham K., Travers M.,
RA Talman J., Tesfaye S., Theodore J., Topham K., Wang S., Wilson B.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilmson B.,
RA Winder E., Wymann D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RT Substratum graminearum genome sequence."
RU "Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AACM0100010; EAA69670.1; -; Genomic_DNA.
KM Hypothetical protein.
SQ SEQUENCE 77 AA; 7980 MW; 72702C4A47662E61 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 77;
Best Local Similarity 71.4%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 CXPXTGC 10
DB 51 CXPXTGC 57

RESULT 35
067W06_ORYZA PRELIMINARY; PRT; 269 AA.
ID 067W06_ORYZA PRELIMINARY; PRT; 269 AA.
AC 067W06;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative senescence-associated protein 5.
GN Name=OSJNB0085J13.14;
OS Oryza sativa (japonica cultivar-group). Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC
RT clone:OSJNB0085J13.1";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; A0003565; B0037413.1; -; Genomic_DNA.
DR Gramene; 067W06; -;
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR000301; Transmem_4.
DR Pfam; PF00335; Tetranspanin; I.
DR PRINTS; PR00259; TMFOUR.
DR PROSITE; PS00421; TM4.1; 1.
SQ SEQUENCE 269 AA; 25488 MW; 8B8EAF766DEB13 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 269;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 172 CXPXTGC 178

RESULT 36
05NVV4_9ARCH PRELIMINARY; PRT; 330 AA.
ID 05NVV4_9ARCH PRELIMINARY; PRT; 330 AA.
AC 05NVV4;

DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Predicted lipase synthase.
GN ORFName=orf120;
OS uncultured archaeon.
OC Archaea; environmental samples.
NCBI_TaxID=115547;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Eikel C., Kemnitz D., Kube M., Rieke P., Chin K.-J., Dedysh S.,
RA Reinhardt R., Conrad R., Liesack W.;
RT "Retrieval of first genome data for rice cluster I methanogens by a
RT combination of cultivation and molecular techniques."
RL FEMS Microbiol. Ecol. 53:187-204(2005).
DR EMBL; CR62685; CAH04852.1; -; Genomic_DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR InterPro; IPR006638; EIP3/MAB/NiIB.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF04055; Radical SAM.
DR SMART; SM00729; EIP3; I.
SQ SEQUENCE 330 AA; 35990 MW; 5090ECC290E75C34 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 330;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 CXPXTGC 10
DB 316 CXPXTGC 322

RESULT 37
07XV21_ORYZA PRELIMINARY; PRT; 464 AA.
ID 07XV21_ORYZA PRELIMINARY; PRT; 464 AA.
AC 07XV21;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE OSJNB0064H22.11 protein.
GN Name=OSJNB0064H22.11;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Meng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu W., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Yu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang U., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4."
RL Nature 420:316-320(2002).
DR EMBL; AL060448; CAD40872.2; -; Genomic_DNA.
DR Gramene; 07XV21; -;

DR GO; GO:0004194; F:pepsin A activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR001461; Peptidase_A1.
DR Pfam; PF00026; Asp; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN 1.
SQ SEQUENCE 464 AA; 45354 MW; 43B1443D3B732501 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 464;

Best Local Similarity 71.4%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 117 CXPXTGC 123

RESULT 38
Q9LEZ5 ARATH

ID Q9LEZ5_ARATH PRELIMINARY; PRT; 1058 AA.

AC Q9LEZ5;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DE Hypothetical protein T22D6_50.

GN Name=T22D6_50;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC euroside II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OC NCBI_TaxID=3702;

OX NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OX NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OX NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=10090;

OC NCBI_TaxID=1

```

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE P110035 protein.
GN OrderedLocustNames=p110035;
OS Listeria innocua.
OC Plasmid p11100.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaeser P., Frangoul L., Buchrieser C., Ruenik C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Chabrie A., Checouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussutget O.,
RA Entlan K.-D., Fshih H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,
RA Jones L.-M., Kaerst U., Krett J., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordleier G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL, AL592102; CAC42033.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0016564; F:transcriptional repressor activity; IEA.
DR GO; GO:0016481; P:negative regulation of transcription; IEA.
DR GO; GO:0046685; P:response to arsenic; IEA.
DR InterPro; IPR010712; Arsd.
DR InterPro; IPR012335; ThioRedoxin_fold.
DR Pfam; PF06953; Arsd; 1.
KW Complete proteome; Plasmid.
SQ SEQUENCE 123 AA; 13164 MW; 5AB2979019CFF14F CRC64;

Query Match 72.7%; Score 32; DB 2; Length 123;
Best Local Similarity 57.1%; Pred. No. 1.4e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
Db 116 CSPSSGC 122

RESULT 42
OSTN00 ANOGA PRELIMINARY; PRT; 222 AA.
AC OSTN00.
DT 01-FEB-2005 (TReMBLrel. 29, Created)
DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)
DE ENSANGP00000028786 (Fragment).
GN ORFNames=ENSANG0000002818;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST.
RG The Anopheles gambiae Sequence Committee;
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST.
RG The Anopheles gambiae Sequence Committee;
RT Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

```

```

CC preliminary data.
DR EMBL; AAB01008986; EAL38762.1; -; Genomic_DNA.
DT NON_TER 1
SQ SEQUENCE 222 AA; 21986 MW; 3628579D88CCE889 CRC64;

Query Match 72.7%; Score 32; DB 2; Length 222;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
Db 78 CSPSSGC 84

RESULT 43
OSTM29 ANOGA PRELIMINARY; PRT; 231 AA.
AC OSTM29.
DT 01-FEB-2005 (TReMBLrel. 29, Created)
DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)
DE ENSANGP00000026276 (Fragment).
GN ORFNames=ENSANG0000002818;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST.
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST.
RG The Anopheles gambiae Sequence Committee;
RT Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AAB01008986; EAL38763.1; -; Genomic_DNA.
DT NON_TER 1
FT NON_TER 231
SQ SEQUENCE 231 AA; 23375 MW; 888B95ACB094A91B CRC64;

Query Match 72.7%; Score 32; DB 2; Length 231;
Best Local Similarity 57.1%; Pred. No. 2.5e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
Db 86 CSPSSGC 92

RESULT 44
O6Z8K1 ORYSA PRELIMINARY; PRT; 520 AA.
AC O6Z8K1.
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Aspartyl protease-like.
GN Name=P0431B06.25;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;

```

```

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004774; BAD15642.1; -; Genomic_DNA.
DR HSSP; P56272; 1AMS.
DR Gramene; O628K1; -.
DR GO; GO:0004194; F:pepsin A activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001461; Peptidase A1.
DR InterPro; IPR001969; Pept_Asp_AS.
DR Pfam; PF00026; Asp_1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
KW Protease.
SQ SEQUENCE 520 AA; 56271 MW; 89468ACBA8C0CA1C CRC64;

Query Match
Best Local Similarity 72.7%; Score 32; DB 2; Length 520;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10
Db 170 CSPASGC 176

RESULT 45
Q9F126 ARATH PRELIMINARY; PRT; 1152 AA.
AC Q9F126_
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MDN11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxId=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99397451; PubMed=10470850;
RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
RT Sequence features of the regions of 1,011,550 bp covered by seventeen
RT P1 and TAC clones."
RL DNA Res. 6:183-195 (1999).
DR EMBL; AB017064; BAB1076.1; -; Genomic_DNA.
DR EMBL; AB023039; BAB1076.1; JOINED; Genomic_DNA.
DR GO; GO:0005488; F:binding; IEA.
DR InterPro; IPR011989; ARM-1like.
SQ SEQUENCE 1152 AA; 126806 MW; FA379C6BF94CA04 CRC64;

Query Match
Best Local Similarity 57.1%; Score 32; DB 2; Length 1152;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10
Db 475 CSPTSGC 481

RESULT 46
Q9WU17 MESAU PRELIMINARY; PRT; 162 AA.
AC Q9WU17_
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Receptor tyrosine kinase (Fragment).
GN Name=Tyrc;
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

```

```

OC Muridae; Cricetinae; Mesocricetus.
OX NCBI_TaxId=10036;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jensen H.T., Lehman M.N., Stevens P.J.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF104241; MAD21012.1; -; mRNA.
DR HSSP; O62838; 1LUF.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0004714; F:transferrase activity; IEA.
DR GO; GO:0004714; F:transmembrane receptor protein tyrosine kin. . .; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002011; ReceptTyrcKinII.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR ProDom; PD000001; Prot_kinase_1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
KW Kinase; Receptor.
FT NON_TER 1
FT TER 162
SQ SEQUENCE 162 AA; 18209 MW; 29DF9E933BECB243 CRC64;

Query Match
Best Local Similarity 70.5%; Score 31; DB 2; Length 162;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10
Db 96 CSPSAGC 102

RESULT 47
Q8T415 DROME PRELIMINARY; PRT; 166 AA.
AC Q8T415_
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE AT044499 (CG31128-PA).
GN ORFName=CG31128;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Stapleton M., Broksstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Padleb J., Paragae V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blake J.R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Bendale J., Bayraktaroglu U., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Broksstein P., Brodtier P.,

```

RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
 RA de Fabios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
 RT *melanogaster* euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 [4]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RT a genomic perspective.";
 RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
 [5]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Mistra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,
 RA Smith C.D., Tuzy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 [6]
 RP NUCLEOTIDE SEQUENCE.
 RG Berkeley *Drosophila* Genome Project;
 RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
 RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
 RA Yu C., Rubin G.;
 RT "Drosophila melanogaster release 4 sequence.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 [7]
 RP NUCLEOTIDE SEQUENCE.
 RG FlyBase;

RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY084097; AA189835.1; -; mRNA.
 DR EMBL; AE003748; AA13997.1; -; Genomic_DNA.
 DR Ensemble; CG31128; *Drosophila melanogaster*.
 DR FlyBase; FBgn0051128; CG31128.
 SQ SEQUENCE 166 AA; 18434 MW; 054DC42BDE653B76 CRC64;
 Query Match 70.5%; Score 31; DB 2; Length 166;
 Best Local Similarity 57.1%; Pred. No. 2.9e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 4 CXPXTGC 10
 DB 142 CPTAGC 148
 RESULT 48
 ID 04H68_9DEIO PRELIMINARY; PRT; 183 AA.
 AC 04H68;
 DT 13-SEP-2005 (TRENBLREL. 31, Created)
 DT 13-SEP-2005 (TRENBLREL. 31, Last sequence update)
 DE Hypothetical protein.
 GN ORFNames=DgeODRAFT_2849;
 OS Deinococcus geothermalis DSM 11300.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 OX NCBI_TaxID=319795;
 RX NUCLEOTIDE SEQUENCE.
 RP STRAIN=DSM 11300;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Deter C., Glavina T.,
 RA Hammon N., Istrati S., Pittluck S., Richardson P.;
 RT "Sequencing of the draft genome assembly of *Deinococcus geothermalis*
 RT DSM 11300.";
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 [2]
 RP NUCLEOTIDE SEQUENCE.
 RG STRAIN=DSM 11300;
 RT US DOE Joint Genome Institute (JGI-ORNL);
 RA Lartier F., Land M.;
 RT "Annotation of the draft genome assembly of *Deinococcus geothermalis*
 RT DSM 11300.";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAHE0100014; EAL81853.1; -; Genomic_DNA.
 KM Hypothetical protein.
 SQ SEQUENCE 183 AA; 19839 MW; C5F6864281158966 CRC64;
 Query Match 70.5%; Score 31; DB 2; Length 183;
 Best Local Similarity 57.1%; Pred. No. 3.1e+02;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 4 CXPXTGC 10
 DB 167 CPTAGC 173
 RESULT 49
 ID 04S7H4_TETNG PRELIMINARY; PRT; 229 AA.
 AC 04S7H4;
 DT 13-SEP-2005 (TRENBLREL. 31, Created)
 DT 13-SEP-2005 (TRENBLREL. 31, Last sequence update)
 DT 13-SEP-2005 (TRENBLREL. 31, Last annotation update)
 DE Chromosome 13 SCAR14715, whole genome shotgun sequence.
 GN ORFNames=GSTENG00022782001;
 OS Tetraodon nigroviridis (green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 OX NCBI_TaxID=99883;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
 Mauceli B., Bouneau L., Fischer C., Ozouf-Costaz C., Benoit A.,
 Nicard S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Segurens B.,
 Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 Anthouard V., Jubin C., Castel V., Katinka M., Vacherie B.,
 Biemont C., Skallil Z., Cattolico L., Poulain J., De Bernardis V.,
 Craud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
 Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
 Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
 Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 Laudet V., Schachter V., Queller F., Saurin W., Scarpelli C.,
 Wincker P., Lander E.S., Weissbach J., Roest Crollius H.,
 RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 the early vertebrate proto-karyotype.",
 RT Nature 431:946-957(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope; Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL; CAE01014715; CAG03408.1; -; Genomic DNA.
 SQ SEQUENCE 229 AA; 24963 MW; 4206FE2E3CCE9F7F CRC64;

Query Match 70.5%; Score 31; DB 2; Length 229;
 Best Local Similarity 57.1%; Pred. No. 3.9e+02;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CKPXTGC 10
 | | | |
 Db 113 CSPGSGC 119

RESULT 50
 Q93KV9 STRVR PRELIMINARY; PRT; 239 AA.
 ID Q93KV9;
 AC Q93KV9;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative methyltransferase.
 GN Name=aviG6;
 OS Streptomyces viridochromogenes.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1938;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA STRAIN=The57;
 RX MEDLINE=2103144; PubMed=11410376; DOI=10.1016/S1074-5521(01)00040-0;
 RA Weinbauer G., Muhlenweg A., Trefter A., Hofmeister D., Sussmuth R.D.,
 Jung G., Weizel K., Vente A., Gireser U., Bechtold A.,
 RT "Biosynthesis of the orthosomycin antibiotic avilamycin A: deductions
 from the molecular analysis of the avl biosynthetic gene cluster of
 RT Streptomyces viridochromogenes Tu57 and production of new
 RT antibiotics";
 RT Chem. Biol. 8:569-581(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA STRAIN=The57;
 RA Mosbacher T., Weinbauer G., Bechtold A., Schulz G.B.,
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AR333038; AAK83186.1; -; Genomic DNA.
 DR GO; GO:0008168; F:methyltransferase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR006342; F:KDM_mtfase.

DR TIGRFAms; TIGR01444; F:KDM_fam; 1.
 KW Methyltransferase; Transferase.
 SQ SEQUENCE 239 AA; 26314 MW; A2FA05654E31A60 CRC64;

Query Match 70.5%; Score 31; DB 2; Length 239;
 Best Local Similarity 57.1%; Pred. No. 4e+02;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CKPXTGC 10
 | | | |
 Db 22 CAPDSC 28

Search completed: January 4, 2006, 16:07:54
 Job time : 125.896 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 : Search time 13.4522 Seconds
(without alignments)
92.983 Million cell updates/sec

Title: US-09-932-322-1

Perfect score: 44
Sequence: 1 XXXXXPTGXXXX 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	81.8	601	2 A27020	DIF-induced preste
2	34	77.3	145	2 S07983	phospholipase A2 h
3	34	77.3	312	2 T05348	hypothetical prote
4	33	75.0	1058	2 T50496	hypothetical prote
5	32	72.7	64	2 P96006	hypothetical prote
6	31	70.5	482	2 T48397	S-receptor kinase-
7	31	70.5	643	2 T25473	hypothetical prote
8	30	68.2	129	2 A24255	chorion class A pr
9	30	68.2	132	2 B24255	chorion class A pr
10	30	68.2	151	2 AB0331	sigma B factor reg
11	30	68.2	199	2 T49486	hypothetical prote
12	30	68.2	200	2 S04926	wound-induced prote
13	30	68.2	201	2 T07729	wound-induced prote
14	30	68.2	211	2 T04098	CBP20 preproprotei
15	30	68.2	211	2 S04927	wound-induced prote
16	30	68.2	222	1 MWVZB4	antithrombin-III h
17	30	68.2	224	2 E84326	hypothetical prote
18	30	68.2	336	2 A83801	ABC transpore (s
19	30	68.2	345	1 MWVZM2	antithrombin-III h
20	30	68.2	406	2 T23898	hypothetical prote
21	30	68.2	430	2 T23899	hypothetical prote
22	30	68.2	551	2 S01793	arylsulfatase (EC
23	30	68.2	551	2 S07089	arylsulfatase (EC
24	30	68.2	648	2 T23864	hypothetical prote
25	30	68.2	1722	2 E89753	protein P11C7.4 li
26	29	65.9	46	2 H71262	hypothetical prote
27	29	65.9	119	2 B45937	early chorion prote
28	29	65.9	119	2 S24292	chorion protein -
29	29	65.9	119	2 S24294	chorion protein -

30	29	65.9	119	2 S24291	Chorion protein -
31	29	65.9	121	2 S24293	Chorion class CA p
32	29	65.9	147	2 E82523	hypothetical prote
33	29	65.9	163	2 T33130	hypothetical prote
34	29	65.9	245	2 F84680	hypothetical prote
35	29	65.9	333	2 P90172	hypothetical prote
36	29	65.9	348	2 T35248	probable oxidoredu
37	29	65.9	375	2 A83636	hypothetical prote
38	29	65.9	379	2 S14885	hypothetical prote
39	29	65.9	449	2 E96676	hypothetical prote
40	29	65.9	507	2 T23375	hypothetical prote
41	29	65.9	513	2 D88991	protein apx-1 (imp
42	29	65.9	685	2 JC7570	Delta-4 protein -
43	29	65.9	686	2 JC7569	hypothetical prote
44	29	65.9	753	2 T28787	hypothetical prote
45	29	65.9	843	1 A27340	complement C7 prec
46	29	65.9	1101	2 T16840	hypothetical prote
47	29	65.9	1146	2 T02766	adhesin W1-1 - Aje
48	29	65.9	1743	2 T26859	hypothetical prote
49	29	65.9	3635	2 T10053	laminin alpha 5 ch
50	28	63.6	74	2 S25773	testis-specific pr
51	28	63.6	95	2 A13171	conserved hypothet
52	28	63.6	103	2 G84741	hypothetical prote
53	28	63.6	105	2 S23061	hypothetical prote
54	28	63.6	110	2 S16496	hypothetical prote
55	28	63.6	154	2 G72564	cysteine proteinas
56	28	63.6	166	2 B44938	hypothetical prote
57	28	63.6	178	2 S76878	hypothetical prote
58	28	63.6	207	2 G69381	flagellin (flaB1-2
59	28	63.6	207	2 B71943	hypothetical prote
60	28	63.6	218	2 T33545	hypothetical prote
61	28	63.6	237	2 I47031	insulin-like growt
62	28	63.6	248	2 D86012	insulin-like growt
63	28	63.6	248	2 D91166	probable membrane
64	28	63.6	254	2 I46603	insulin-like growt
65	28	63.6	254	2 I48599	insulin-like growt
66	28	63.6	254	2 JC1464	insulin-like growt
67	28	63.6	258	1 B37252	insulin-like growt
68	28	63.6	258	2 A45403	insulin-like growt
69	28	63.6	316	2 AG0121	probable sideropho
70	28	63.6	331	2 AC0842	glycine betaine-bi
71	28	63.6	340	1 WMBEL1	latency-related pr
72	28	63.6	449	1 L1Pp	retinoic acid rece
73	28	63.6	453	2 I50674	ankyrin-related pr
74	28	63.6	491	2 B87452	arylsulfatase (EC
75	28	63.6	567	2 A37362	urase (EC 3.5.1.5
76	28	63.6	569	2 T50711	hypothetical prote
77	28	63.6	628	2 S50524	hypothetical prote
78	28	63.6	744	2 A43353	ascites sialoglyco
79	28	63.6	873	2 H96503	protein P9C16.17 l
80	28	63.6	910	2 A43721	androgen receptor
81	28	63.6	911	2 B34721	androgen receptor
82	28	63.6	919	2 A35248	androgen receptor
83	28	63.6	980	2 T00045	cellulodextrin phosp
84	28	63.6	1035	2 S18512	protein P9C16.13 l
85	28	63.6	1172	2 P96503	protein P9C16.13 l
86	28	63.6	1220	2 A56136	jaaged protein pre
87	28	63.6	1332	2 T23024	hypothetical prote
88	28	63.6	1522	2 H88380	protein T22F7.3 li
89	28	63.6	1620	2 T27283	hypothetical prote
90	28	63.6	1717	1 A45558	epidermal growth f
91	28	63.6	2395	1 S50820	surface protein ty
92	28	63.6	2533	2 T28675	alpha-51D immobili
93	28	63.6	2533	2 T28674	alpha-51D-immobili
94	28	63.6	2543	2 T31687	surface antigen - P
95	28	63.6	2543	2 S09118	G surface protein
96	28	63.6	2704	2 A23475	G surface protein
97	27	61.4	2718	2 A44379	alpha-conotoxin SI
98	27	61.4	46	1 KECX	crimbin - Alysini
99	27	61.4	45	1 A40186	ribosomal protein
100	27	61.4	54	2 F82845	hypothetical prote
101	27	61.4	75	2 S17156	hypothetical prote
102	27	61.4	87	2 S41306	metallothionein -

Chorion protein -
Chorion class CA p
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
probable oxidoredu
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
protein apx-1 (imp
Delta-4 protein -
hypothetical prote
complement C7 prec
hypothetical prote
adhesin W1-1 - Aje
hypothetical prote
laminin alpha 5 ch
testis-specific pr
conserved hypothet
hypothetical prote
Chorion protein -
hypothetical prote
hypothetical prote
cysteine proteinas
hypothetical prote
flagellin (flaB1-2
hypothetical prote
hypothetical prote
insulin-like growt
insulin-like growt
insulin-like growt
insulin-like growt
probable membrane
insulin-like growt
insulin-like growt
insulin-like growt
insulin-like growt
probable sideropho
glycine betaine-bi
latency-related pr
retinoic acid rece
ankyrin-related pr
arylsulfatase (EC
urase (EC 3.5.1.5
hypothetical prote
hypothetical prote
ascites sialoglyco
protein P9C16.17 l
androgen receptor
androgen receptor
androgen receptor
cellulodextrin phosp
protein P9C16.13 l
jaaged protein pre
hypothetical prote
protein T22F7.3 li
hypothetical prote
epidermal growth f
surface protein ty
alpha-51D immobili
alpha-51D-immobili
surface antigen - P
G surface protein
G surface protein
alpha-conotoxin SI
crimbin - Alysini
ribosomal protein
hypothetical prote
hypothetical prote
metallothionein -
hypothetical prote

103	27	61.4	89	2	T42019	Ser/Thr protein ph
104	27	61.4	99	2	D70582	probable PR protei
105	27	61.4	118	1	PSUT1B	phospholipase A2 (
106	27	61.4	118	2	S52549	thionin variant Th
107	27	61.4	120	1	W4WL42	B4 protein - human
108	27	61.4	120	2	E72490	hypothetical prote
109	27	61.4	121	2	F81973	probable ribonucle
110	27	61.4	121	2	H81027	ribonuclease P pro
111	27	61.4	123	2	T19391	hypothetical prote
112	27	61.4	125	2	S52548	thionin variant Th
113	27	61.4	125	2	S52550	thionin variant Th
114	27	61.4	128	2	A72678	hypothetical prote
115	27	61.4	130	2	T08584	hypothetical prote
116	27	61.4	134	2	AD2592	cytidine deaminase
117	27	61.4	135	1	HSX132	histone H3.2 - Afr
118	27	61.4	136	1	S52545	thionin variant Th
119	27	61.4	145	2	F97374	cytidine deaminase
120	27	61.4	150	2	T49563	hypothetical prote
121	27	61.4	155	2	A10545	hypothetical prote
122	27	61.4	173	2	T11112	NADH2 dehydrogena
123	27	61.4	190	2	T05492	thiamatin homolog
124	27	61.4	219	2	T51074	hypothetical prote
125	27	61.4	220	2	H81434	superoxide dismuta
126	27	61.4	231	2	S55078	hypothetical prote
127	27	61.4	234	2	T36162	probable integral
128	27	61.4	237	2	S08073	cyclic nucleotide
129	27	61.4	237	2	T25152	hypothetical prote
130	27	61.4	239	2	A25317	gag polyprotein -
131	27	61.4	246	2	A64326	hypothetical prote
132	27	61.4	248	2	S30194	ribosomal protein
133	27	61.4	249	1	R3HU6	ribosomal protein
134	27	61.4	249	1	R3MS6	ribosomal protein
135	27	61.4	249	1	R3RMS6	ribosomal protein
136	27	61.4	249	2	UC4145	ribosomal protein
137	27	61.4	249	2	S41468	ribosomal protein
138	27	61.4	249	2	T50003	40S ribosomal prote
139	27	61.4	250	2	T05100	ribosomal protein
140	27	61.4	251	2	T04334	erythromycin resis
141	27	61.4	253	2	S09215	p19 protein - avia
142	27	61.4	254	2	S35743	lymphocyte activat
143	27	61.4	255	2	I38426	T-cell antigen 4-1
144	27	61.4	259	2	B32393	hypothetical prote
145	27	61.4	259	2	T21011	cyclo protein limp
146	27	61.4	265	2	C87679	probable senescen
147	27	61.4	270	2	B84578	gene ox40 protein
148	27	61.4	271	2	S12783	B-cell activation
149	27	61.4	272	2	I48700	Ox40 homolog - hum
150	27	61.4	277	2	A60771	B cell-associated
151	27	61.4	277	2	I37552	extracellular matr
152	27	61.4	305	2	A46476	hypothetical prote
153	27	61.4	313	2	S44208	another-specific pr
154	27	61.4	318	2	T29479	t2 protein - rabbl
155	27	61.4	321	2	S13550	probable transcrip
156	27	61.4	325	2	B43692	glycine betaine/pr
157	27	61.4	326	1	GOVZML	hypothetical prote
158	27	61.4	336	1	C95902	glycine betaine/pr
159	27	61.4	330	1	B1RCGP	glycine betaine/pr
160	27	61.4	330	2	T25169	hypothetical prote
161	27	61.4	330	2	A85916	glycine betaine/pr
162	27	61.4	330	2	F91071	fructose-bisphosph
163	27	61.4	334	2	AG0322	hypothetical prote
164	27	61.4	337	2	A10427	lipase homolog T29
165	27	61.4	344	2	T17266	probable myo-inosi
166	27	61.4	381	2	T06696	hypothetical prote
167	27	61.4	388	2	T36191	hypothetical prote
168	27	61.4	392	2	H84679	GTP-binding protei
169	27	61.4	396	2	F84326	exo-alpha-stalidas
170	27	61.4	418	2	JC7588	hypothetical prote
171	27	61.4	432	2	T51468	hypothetical prote
172	27	61.4	436	2	AE3341	cobG protein (EC 1
173	27	61.4	487	2	UC7328	amino acid transpo
174	27	61.4	512	2	S28267	phosphorin I prec
175	27	61.4	528	2	B75310	conserved hypoche
176	27	61.4	533	2	H71492	probable hsp-60 -
177	27	61.4	534	2	B81654	60 kDa chaperonin
178	27	61.4	539	2	T31901	hypothetical prote
179	27	61.4	561	1	SYRPM	malate synthase (E
180	27	61.4	562	2	T48413	malate synthase-ii
181	27	61.4	563	2	C95874	conserved hypoche
182	27	61.4	566	2	T07690	malate synthase (E
183	27	61.4	566	2	S15387	malate synthase (E
184	27	61.4	567	1	SYCSM2	malate synthase (E
185	27	61.4	567	1	SYCSM2	malate synthase (E
186	27	61.4	568	1	SYKMA	protein f54Cl.1 (i
187	27	61.4	581	2	B87768	probable acylamino
188	27	61.4	591	2	H72474	guanosine-3',5'-bi
189	27	61.4	607	2	F70124	gag polyprotein -
190	27	61.4	701	2	S35430	Trg protein - frui
191	27	61.4	720	1	A55160	S-receptor kinase
192	27	61.4	824	1	S50767	ribonucleoside-dip
193	27	61.4	826	1	Q0BE11	ubiquitin ligase N
194	27	61.4	887	2	S70642	pullulanase - spin
195	27	61.4	964	2	S51324	ribonucleotide red
196	27	61.4	987	2	D97029	prestalk protein p
197	27	61.4	1046	2	A26838	hemocytin - silkw
198	27	61.4	3133	2	S52093	tenascin-X precurs
199	27	61.4	3566	1	A40701	MEG1 protein - ra
200	27	61.4	4351	2	T00252	cardioactive pepti
201	26	59.1	9	2	A26363	potassium channel
202	26	59.1	9	2	S27233	hypothetical prote
203	26	59.1	35	2	S69599	hypothetical prote
204	26	59.1	45	2	D82788	hypothetical prote
205	26	59.1	53	2	E82628	testis-specific pr
206	26	59.1	56	1	WTRF	testis-specific pr
207	26	59.1	63	2	S25772	somatotropin precu
208	26	59.1	77	2	PC1017	T-cell surface gly
209	26	59.1	80	2	C49050	hypothetical prote
210	26	59.1	89	2	S75008	transposase s81192
211	26	59.1	93	2	S75008	phospholipase A2 (
212	26	59.1	118	1	PSL73E	phospholipase A2 (
213	26	59.1	118	2	C34860	phospholipase A2 (
214	26	59.1	118	2	F34860	phospholipase A2 h
215	26	59.1	119	1	PSNOA1	alpha-amylose inh
216	26	59.1	121	2	S16920	alpha-amylose inh
217	26	59.1	123	1	WMTA	hypothetical prote
218	26	59.1	123	2	G75336	conserved hypoche
219	26	59.1	126	1	B64025	T-cell activation
220	26	59.1	129	2	A27562	hypothetical prote
221	26	59.1	129	2	AB2728	phospholipase A2 (
222	26	59.1	134	2	A59055	phospholipase A2 (
223	26	59.1	134	2	B59055	hypothetical prote
224	26	59.1	134	2	E70882	phospholipase A2 (
225	26	59.1	145	2	S01390	phospholipase A2 (
226	26	59.1	145	2	S15404	NGF5 secreted spl
227	26	59.1	151	2	T06517	alpha-amylose inh
228	26	59.1	154	2	AB2509	hypothetical prote
229	26	59.1	160	2	S76328	hypothetical prote
230	26	59.1	162	1	PSHBA	phospholipase A2 (
231	26	59.1	170	2	E82145	conserved hypoche
232	26	59.1	171	2	S57894	laminin - Hydra vu
233	26	59.1	172	2	S72644	pilx protein precu
234	26	59.1	176	2	T02217	NBS-IIR type resis
235	26	59.1	181	2	T03720	4-coumarate-CoA 11
236	26	59.1	182	2	T02091	ribosomal protein
237	26	59.1	187	2	G70688	probable lipv prot
238	26	59.1	188	2	H38192	chitin synthase (E
239	26	59.1	188	2	C83076	type 4 fibrial b1
240	26	59.1	195	2	A45067	laminin B1 chain v
241	26	59.1	199	2	UE0351	OX40 ligand protei
242	26	59.1	202	2	T48709	hypothetical prote
243	26	59.1	204	2	A54560	TPA-induced protei
244	26	59.1	205	1	P8XRBH	nonstructural prot
245	26	59.1	205	2	T04370	perit protein - ba
246	26	59.1	207	2	A64655	hypothetical prote
247	26	59.1	207	2	AC3555	hypothetical prote
248	26	59.1	212	2	A72648	hypothetical prote

249	26	59.1	217	1	VCfMS	coat protein - alf	322	26	59.1	383	2	F86411	purtoein FlK23.16
250	26	59.1	218	1	VCfV2S	coat protein - alf	323	26	59.1	384	2	S68410	lipase Arab-1 - Ar
251	26	59.1	219	2	B83396	probable COA trans	324	26	59.1	387	2	B84721	hypothetical prote
252	26	59.1	220	2	S47166	superoxide dismuta	325	26	59.1	390	2	F69452	probable aspartate
253	26	59.1	221	1	VCfM42	coat protein - alf	326	26	59.1	390	2	B86411	protein FlK23.18 (
254	26	59.1	221	1	VCfMYS	coat protein - alf	327	26	59.1	390	2	G82844	cyteine synthase
255	26	59.1	225	2	T20325	hypothetical prote	328	26	59.1	393	2	T51108	dehydroatase [lmpo
256	26	59.1	228	2	F75390	hypothetical prote	329	26	59.1	394	2	A84672	probable lipase (I
257	26	59.1	229	2	S66342	ribonuclease II (B	330	26	59.1	397	2	T46223	hypothetical prote
258	26	59.1	230	2	S53506	ribonuclease LB (B	331	26	59.1	400	2	C75336	serine proteinase,
259	26	59.1	232	2	T33704	hypothetical prote	332	26	59.1	403	2	UE0114	zinc-finger protei
260	26	59.1	237	2	S53507	starvation-induced	333	26	59.1	404	2	T06761	hypothetical prote
261	26	59.1	237	2	T05973	permatin homolog P	334	26	59.1	407	2	AE3615	glycosyl transfera
262	26	59.1	247	2	AD3649	transcription regu	335	26	59.1	408	2	G70522	probable transposa
263	26	59.1	249	2	T35589	probable secreted	336	26	59.1	410	2	T47926	hypothetical prote
264	26	59.1	253	2	S73473	DNA polymerase III	337	26	59.1	410	2	T51212	related to integra
265	26	59.1	253	2	T25768	hypothetical prote	338	26	59.1	412	2	T38742	hypothetical prote
266	26	59.1	258	2	A29088	Smh class II histo	339	26	59.1	420	2	C69532	conserved hypotet
267	26	59.1	266	2	AG3000	1-acyl-sn-glycerol	340	26	59.1	422	2	D86339	hypothetical protea
268	26	59.1	266	2	B98283	hypothetical prote	341	26	59.1	423	2	A41207	collagen 13, nonfi
269	26	59.1	269	2	T04095	hypothetical prote	342	26	59.1	426	2	B84531	hypothetical prote
270	26	59.1	270	2	F69362	ribonuclease S hom	343	26	59.1	427	2	D84540	hypothetical prote
271	26	59.1	271	2	C86251	conserved hypotet	344	26	59.1	427	2	T48159	hypothetical prote
272	26	59.1	274	2	F85741	hypothetical prote	345	26	59.1	430	2	B64129	adenosylmethionine
273	26	59.1	276	2	C84471	probable U4/U6\sm	346	26	59.1	433	2	F84215	aminopeptidase [lm
274	26	59.1	277	2	A75112	hypothetical prote	347	26	59.1	435	2	I54182	tumor necrosis fac
275	26	59.1	277	2	G71021	hypothetical prote	348	26	59.1	446	2	T47744	hypothetical prote
276	26	59.1	281	2	AD2755	conserved hypotet	349	26	59.1	446	2	C75087	hydroxymethylpyrim
277	26	59.1	281	2	AB3367	maze protein (lmpo	350	26	59.1	446	2	B71057	probable thiamin b
278	26	59.1	284	2	E70506	hypothetical prote	351	26	59.1	455	2	I54353	tetracycline trans
279	26	59.1	285	2	T47500	hypothetical prote	352	26	59.1	463	2	A83762	RNA methyltransfer
280	26	59.1	285	2	E75507	hypothetical prote	353	26	59.1	464	2	S75831	hypothetical prote
281	26	59.1	285	2	H95969	hypothetical prote	354	26	59.1	466	2	F69805	RNA methyltransfer
282	26	59.1	286	2	B84807	probable RNA-bind	355	26	59.1	468	1	B66114	glycoprotein gp13
283	26	59.1	292	2	C97536	maze protein (Af22	356	26	59.1	468	1	VG8EBH	glycoprotein gp13
284	26	59.1	295	2	T15193	sporulation transc	357	26	59.1	481	2	T48404	3-deoxy-D-manno-oc
285	26	59.1	297	2	T18960	hypothetical prote	358	26	59.1	495	2	F43496	envelope protein E
286	26	59.1	297	2	S06267	surface antigen H	359	26	59.1	495	2	B43496	envelope protein E
287	26	59.1	298	2	C82318	chromosome initiat	360	26	59.1	495	2	C43496	envelope protein E
288	26	59.1	301	2	H70644	hypothetical prote	361	26	59.1	495	2	B43496	envelope protein E
289	26	59.1	305	2	A97158	stage III sporulat	362	26	59.1	495	2	G43496	envelope protein E
290	26	59.1	306	2	F64185	D-alanine-D-alalan	363	26	59.1	495	2	S06740	envelope protein E
291	26	59.1	306	2	H86695	ABC transporter pe	364	26	59.1	495	2	H43496	envelope protein E
292	26	59.1	308	2	T35835	hypothetical prote	365	26	59.1	495	2	S49137	genome polyprotein
293	26	59.1	310	2	F64146	hypothetical prote	366	26	59.1	495	2	S06741	envelope protein E
294	26	59.1	312	2	A61183	hypothetical prote	367	26	59.1	495	2	S11482	envelope glycoprot
295	26	59.1	314	2	AB0144	conserved hypotet	368	26	59.1	495	2	S06747	envelope protein E
296	26	59.1	315	2	T18956	hypothetical prote	369	26	59.1	496	2	S51668	tyrosine kinase -
297	26	59.1	319	2	AB0641	pseudouridyate sy	370	26	59.1	499	2	S74224	aldehyde dehydroge
298	26	59.1	319	2	C64852	probable pseudouri	371	26	59.1	500	2	JC7668	depeptidyl-peptida
299	26	59.1	319	2	D85671	hypothetical prote	372	26	59.1	503	2	T05347	hypothetical prote
300	26	59.1	319	2	H90811	hypothetical prote	373	26	59.1	505	1	PIWLRB	LI protein - cotro
301	26	59.1	321	1	Q0EC15	hypothetical prote	374	26	59.1	510	1	PWQFA	H+-transporting tw
302	26	59.1	321	2	F91104	hypothetical 34K p	375	26	59.1	510	1	D96741	hypothetical prote
303	26	59.1	321	2	A85950	hypothetical prote	376	26	59.1	511	1	S60287	glucose-6-phosphat
304	26	59.1	326	2	D72689	hypothetical prote	377	26	59.1	512	2	AG3527	D-xylose transport
305	26	59.1	328	2	E85842	probable superoxid	378	26	59.1	515	2	T52610	glucose-6-phosphat
306	26	59.1	332	2	B71165	hypothetical prote	379	26	59.1	519	2	T43756	KTPase subunit 1 (
307	26	59.1	336	2	AG2686	ABC transporter, s	380	26	59.1	522	2	S53842	H+-transporting tw
308	26	59.1	336	2	D97468	hypothetical prote	381	26	59.1	524	2	G73397	MG68 homolog Ro2_
309	26	59.1	338	2	E75576	oxidoreductase - D	382	26	59.1	528	2	S13641	H+-transporting tw
310	26	59.1	338	2	B96976	probable ABC trans	383	26	59.1	530	2	T30505	hypothetical prote
311	26	59.1	341	2	T33949	hypothetical prote	384	26	59.1	531	2	T43551	multidrug resisten
312	26	59.1	341	2	C86461	PI4M2.10 protein -	385	26	59.1	531	2	C83153	conserved hypotet
313	26	59.1	344	2	A75576	oxidoreductase - D	386	26	59.1	532	2	T52442	hypothetical prote
314	26	59.1	353	2	A87469	ubiquinol oxidase	387	26	59.1	545	1	A39827	4-connarate-CoA II
315	26	59.1	359	2	T20575	hypothetical prote	388	26	59.1	545	1	B39827	4-connarate-CoA II
316	26	59.1	360	2	T33835	hypothetical prote	389	26	59.1	551	1	C71606	ATP synthase alpha
317	26	59.1	364	1	PAQXF	fructose-bisphosph	390	26	59.1	554	2	T02445	probable U4/U6 sma
318	26	59.1	367	2	T48048	hypothetical prote	391	26	59.1	555	2	J01405	genome polyprotein
319	26	59.1	367	2	G85362	hypothetical prote	392	26	59.1	555	2	J01404	genome polyprotein
320	26	59.1	378	1	A70571	probable hexosyltr	393	26	59.1	559	2	T09135	hypothetical prote
321	26	59.1	381	2	T45940	hypothetical prote	394	26	59.1	559	2	B64556	para-aminobenzoate

395	26	59.1	560	2	A13216	hypotheical prote
396	26	59.1	561	2	C96837	nodulin-like prote
397	26	59.1	563	2	S33957	coat protein gamma
398	26	59.1	567	2	A71951	p-aminobenzoate sy
399	26	59.1	588	2	T49210	hypotheical prote
400	26	59.1	586	2	T35549	hypotheical prote
401	26	59.1	597	2	G84825	probable CCGH-type
402	26	59.1	603	2	T48154	pyruvate decarboxy
403	26	59.1	621	2	S26691	serine proteinase
404	26	59.1	625	2	C25977	phosphotransferase
405	26	59.1	625	2	T10661	serine/threonine-s
406	26	59.1	642	2	C86152	T7123.2 protein -
407	26	59.1	647	2	T30892	hypotheical prote
408	26	59.1	662	2	D40228	neurexin II-beta p
409	26	59.1	665	2	PS0043	genome polypotein
410	26	59.1	675	2	A35743	creatine kinase (B
411	26	59.1	684	2	G84730	Mutator-like trans
412	26	59.1	716	2	H72763	hypotheical prote
413	26	59.1	727	2	B84847	probable CCGH-type
414	26	59.1	768	2	A42755	P-selectin precurs
415	26	59.1	768	2	I53821	P-selectin - rat
416	26	59.1	775	2	A48644	polypotein(C, E,
417	26	59.1	775	2	A47311	probable membrane-
418	26	59.1	779	2	H71301	inner membrane pro
419	26	59.1	787	2	G81692	probable 60kda inn
420	26	59.1	787	2	E71537	genome polypotein
421	26	59.1	792	2	C32401	genome polypotein
422	26	59.1	792	2	B32401	genome polypotein
423	26	59.1	792	2	A32401	60 kda inner membr
424	26	59.1	795	2	P86529	hypotheical prote
425	26	59.1	795	2	T20609	inner membrane pro
426	26	59.1	795	2	E81575	60 kda inner membr
427	26	59.1	795	2	G72092	squamosa-promoter
428	26	59.1	812	2	T52569	monocyste surfice a
429	26	59.1	826	2	A60385	taase receptor TIR
430	26	59.1	858	2	JC7683	hypotheical prote
431	26	59.1	879	2	C90879	hypotheical prote
432	26	59.1	879	2	G85739	hypotheical prote
433	26	59.1	883	2	A96662	conserved hypotet
434	26	59.1	884	2	E75489	hyphally regulated
435	26	59.1	940	2	S58135	projectin - fruit
436	26	59.1	940	2	A40985	hypotheical prote
437	26	59.1	963	2	T26022	protein-tyrosine k
438	26	59.1	987	2	A54092	probable glnf - My
439	26	59.1	994	2	A70776	BI protein A - black
440	26	59.1	998	1	Q08B81	protein A - flock
441	26	59.1	998	2	S41397	probable two-compo
442	26	59.1	1024	2	G83022	hypotheical prote
443	26	59.1	1024	2	T27631	hypotheical prote
444	26	59.1	1030	2	H88859	protein ZC518.2 [i
445	26	59.1	1035	2	G86342	hypotheical prote
446	26	59.1	1051	2	D82428	chitodextrinase VC
447	26	59.1	1097	1	RNBYL3	DNA-directed DNA p
448	26	59.1	1111	1	T26972	hypotheical prote
449	26	59.1	1127	1	GNWV2	genome polypotein
450	26	59.1	1200	2	T43148	probable protein-t
451	26	59.1	1201	2	H86434	protein F17F8.21 [
452	26	59.1	1226	1	GNWVWP	genome polypotein
453	26	59.1	1240	2	JC5209	insulin receptor s
454	26	59.1	1242	2	JS0670	insulin receptor s
455	26	59.1	1290	2	A57190	ebnerin precursor
456	26	59.1	1375	2	T18861	FAB1 protein homol
457	26	59.1	1428	2	T08852	lustrin A - Califo
458	26	59.1	1438	2	A48216	neurexin III-alpha
459	26	59.1	1466	2	T30566	ATP-binding caset
460	26	59.1	1471	2	B48218	neurexin III-alpha
461	26	59.1	1474	2	D88550	protein ZC84.6 [im
462	26	59.1	1530	1	S01393	DNA-directed RNA p
463	26	59.1	1567	2	S11672	ice nucleation pro
464	26	59.1	1574	2	T13954	MEGF6 protein - ra
465	26	59.1	1578	2	I48216	neurexin III-alpha
466	26	59.1	1592	2	S48933	probable transport
467	26	59.1	1615	2	JC6510	ras-responsive ele
468	26	59.1	1630	2	A53577	aeclies sialoglyco
469	26	59.1	1639	1	NMFPR2	lamnin gamma-1 ch
470	26	59.1	1650	2	S28721	hypotheical prote
471	26	59.1	1715	2	C40228	neurexin II-alpha
472	26	59.1	1732	2	T30836	lysine-specific cy
473	26	59.1	1742	2	S24600	projectin - fruit
474	26	59.1	1766	1	NMHB1	lamnin beta-1 cha
475	26	59.1	1786	1	NMMSB1	lamnin beta-1 cha
476	26	59.1	1790	1	NMFPB1	lamnin beta-2 cha
477	26	59.1	1798	1	S53869	lamnin beta-2 cha
478	26	59.1	1801	1	NMRTS	lamnin beta-2 cha
479	26	59.1	1816	1	S68960	lamnin alpha-4 ch
480	26	59.1	1892	2	T18314	hypotheical prote
481	26	59.1	1905	1	S68235	myosin-light-chain
482	26	59.1	1955	1	AGCH	agrin precursor -
483	26	59.1	1959	1	AGRT	agrin - rat
484	26	59.1	2019	1	UQ1322	tenascin precursor
485	26	59.1	2083	2	T42721	CRP-ductin-alpha p
486	26	59.1	2090	2	T30075	hypotheical prote
487	26	59.1	2101	2	S57245	insulin receptor (
488	26	59.1	2145	2	S61041	glutamate synthase
489	26	59.1	2153	2	T30074	hypotheical prote
490	26	59.1	2344	2	T41590	probable sensor-1i
491	26	59.1	2684	2	A96521	protein F2ID18.22
492	26	59.1	2844	2	S28291	hypotheical prote
493	26	59.1	3163	2	AB0233	yersinabactin bio
494	26	59.1	3163	2	T17440	genome polypotein
495	26	59.1	3388	1	GNWVDP	genome polypotein
496	26	59.1	3390	1	GNWV3	genome polypotein
497	26	59.1	3391	1	GNWV16	genome polypotein
498	26	59.1	3391	1	GNWV26	genome polypotein
499	26	59.1	3391	1	GNWV7A	polypotein - deng
500	26	59.1	3391	2	JS0219	genome polypotein
501	26	59.1	3396	1	A42551	genome polypotein
502	26	59.1	3430	1	GNWVW	lamnin alpha-1 ch
503	26	59.1	3712	2	S18253	genome polypotein
504	26	59.1	3898	1	A44217	genome polypotein
505	26	59.1	3988	1	GNWV8V	genome polypotein
506	26	59.1	5376	2	T42215	zonadhesin - mouse
507	26	59.1	6558	2	T13931	projectin - fruit
508	25	56.8	73	2	C87400	Fold bifunctional
509	25	56.8	73	2	T31219	hypotheical prote
510	25	56.8	82	2	B46308	E3 class 2 protein
511	25	56.8	88	1	TZHQK	killer toxin - yea
512	25	56.8	89	2	B95913	hypotheical expor
513	25	56.8	100	2	A46308	E3 class 1 protein
514	25	56.8	105	2	E41377	hypotheical prote
515	25	56.8	118	1	PSL74E	phospholipase A2 (
516	25	56.8	118	2	T17967	ERV1 protein homol
517	25	56.8	130	1	KRSH3A	keratin high-sulfu
518	25	56.8	131	1	KRSH3A	keratin high-sulfu
519	25	56.8	131	1	KRSH3A	keratin high-sulfu
520	25	56.8	132	1	KRG73U	keratin high-sulfu
521	25	56.8	132	1	KRG73U	keratin high-sulfu
522	25	56.8	132	2	S24302	hypotheical prote
523	25	56.8	133	2	S00385	hypothetical prote
524	25	56.8	144	2	I52649	circadian protein
525	25	56.8	145	2	D75603	trypsin inhibitor
526	25	56.8	145	2	D75603	hypothetical prote
527	25	56.8	161	2	JC4275	pleiotrophic facto
528	25	56.8	161	2	JC4274	pleiotrophic facto
529	25	56.8	161	2	C82703	hypothetical prote
530	25	56.8	162	2	E71368	probable RNA polym
531	25	56.8	174	2	C46308	E3 class 3 protein
532	25	56.8	183	1	I64006	hypothetical prote
533	25	56.8	188	1	JC6547	high sulfur protei
534	25	56.8	195	2	JH0719	omega-conotoxin re
535	25	56.8	202	2	E82396	malate synthase-re
536	25	56.8	204	2	T39940	GTP cyclonhydrolase
537	25	56.8	211	2	S18463	lysosome (EC 3.2.1
538	25	56.8	218	2	H90856	hypothetical prote
539	25	56.8	222	2	B49346	butyrate-acetacac
540	25	56.8	227	2	H86394	ribonuclease (EC 3

541	25	56.8	229	2	JC5518	ribonuclease M5 (B	614	25	56.8	421	2	T00598	hypothetical prote
542	25	56.8	230	2	D81223	conserved hypothet	615	25	56.8	421	2	JC7367	second peroxisomal
543	25	56.8	230	2	C81994	hypothetical prote	616	25	56.8	425	1	A26431	nerve growth facto
544	25	56.8	235	2	B82173	pseudouridine synt	617	25	56.8	425	1	SAVYTL	satellite RNA-enco
545	25	56.8	239	2	S66341	ribonuclease I (Bc	618	25	56.8	426	2	DT1367	conserved hypothet
546	25	56.8	240	2	T04419	alatroxone ribonucle	619	25	56.8	427	1	GQHUN	nerve growth facto
547	25	56.8	243	2	B72757	probable phosphate	620	25	56.8	428	2	AF0302	probable solute-bi
548	25	56.8	248	2	T22755	hypothetical prote	621	25	56.8	428	2	S03767	cellulase (EC 3.2.
549	25	56.8	261	2	H83848	cochlamn synthase	622	25	56.8	433	2	F81940	probable adenosym
550	25	56.8	266	2	A12762	conserved hypothet	623	25	56.8	433	2	D81164	adenosylmethionine
551	25	56.8	271	2	H97543	hypothetical prote	624	25	56.8	440	2	A84387	oligopeptidase ABC
552	25	56.8	274	2	AC3295	heat resistant agg	625	25	56.8	443	2	F89426	protein M162.5 [Im
553	25	56.8	283	2	T02287	hypothetical prote	626	25	56.8	447	2	A47430	gastrin/cholecysto
554	25	56.8	286	2	S44561	hypothetical prote	627	25	56.8	451	2	A45643	tubulin alpha chai
555	25	56.8	291	2	A13419	hypothetical phage	628	25	56.8	453	1	NMIV3	exo-alpha-sialidas
556	25	56.8	298	2	F69515	hypothetical prote	629	25	56.8	453	2	T15374	hypothetical prote
557	25	56.8	299	1	ZZZRNK	3'-phosphodenosin	630	25	56.8	454	1	NMIV	exo-alpha-sialidas
558	25	56.8	299	1	G95932	probable sulfatase	631	25	56.8	455	2	T39537	tubulin alpha-1 ch
559	25	56.8	299	2	D95320	sulfate adenylylitr	632	25	56.8	455	2	A25072	hypothetical prote
560	25	56.8	300	2	A81418	pseudouridylylate sy	633	25	56.8	456	2	A85139	hypothetical prote
561	25	56.8	302	2	C64103	glycine-tRNA ligase	634	25	56.8	461	2	T02706	protein p17p6.19 f
562	25	56.8	302	2	T37514	hypothetical prote	635	25	56.8	461	2	A86435	hypothetical prote
563	25	56.8	303	1	SYRGA	glycine-tRNA ligase	636	25	56.8	462	2	S67569	hypothetical prote
564	25	56.8	303	2	C91184	glycine tRNA synth	637	25	56.8	463	1	S77180	glutamy-tRNA redu
565	25	56.8	303	2	H86030	glycine tRNA synth	638	25	56.8	463	2	S66504	dipeptidyl-peptida
566	25	56.8	303	2	A71557	acetylpliamine am	639	25	56.8	466	2	S36794	beta-1-adrenergic
567	25	56.8	304	1	S74557	glycine-tRNA ligase	640	25	56.8	468	2	B96514	hypothetical prote
568	25	56.8	304	2	A10494	peroxisome assembl	641	25	56.8	469	2	D86144	protein probable U
569	25	56.8	304	2	A45989	hypothetical prote	642	25	56.8	470	1	NMIV07	exo-alpha-sialidas
570	25	56.8	304	2	T24703	ATP sulfurylase	643	25	56.8	470	1	NMIVXL	exo-alpha-sialidas
571	25	56.8	305	2	E83091	ATP sulfurylase sm	644	25	56.8	470	2	S04801	probable membrane
572	25	56.8	305	2	A41812	peroxisome assembl	645	25	56.8	478	2	S51457	3-phytase (EC 3.1.
573	25	56.8	306	2	AB1911	hypothetical prote	646	25	56.8	479	1	JN0715	hypothetical prote
574	25	56.8	306	2	S32834	methylviologen-red	647	25	56.8	481	2	G86144	two-component syst
575	25	56.8	309	2	B87576	hypothetical prote	648	25	56.8	481	2	T35149	GPI-anchor biosynt
576	25	56.8	311	2	B87433	sulfate adenylylate	649	25	56.8	484	1	A46217	gene pIG-A protein
577	25	56.8	314	2	I37383	FAS soluble proteol	650	25	56.8	485	1	A55731	IMP dehydrogenase
578	25	56.8	315	2	E83645	glycyl-tRNA synthet	651	25	56.8	485	2	I52484	IMP dehydrogenase
579	25	56.8	315	2	H83062	sulfate adenylylate	652	25	56.8	487	2	AG0349	IMP dehydrogenase
580	25	56.8	315	2	S19737	hypothetical prote	653	25	56.8	487	2	T23776	hypothetical prote
581	25	56.8	317	2	F82672	ATP sulfurylase, s	654	25	56.8	488	1	D58C1P	IMP dehydrogenase
582	25	56.8	317	2	A12676	orif protein - Aut	655	25	56.8	488	1	H64055	IMP dehydrogenase
583	25	56.8	324	2	E44221	probable modificat	656	25	56.8	488	2	G85894	IMP dehydrogenase
584	25	56.8	327	2	D72503	glycyl-tRNA synthet	657	25	56.8	488	2	B91050	IMP dehydrogenase
585	25	56.8	327	2	G95355	Probable ABC trans	658	25	56.8	490	2	AE0820	hypothetical prote
586	25	56.8	330	2	G82373	glycyl-tRNA synthet	659	25	56.8	495	2	B71360	hypothetical prote
587	25	56.8	331	2	AB3471	sulfate adenylylitr	660	25	56.8	499	1	CPBHS	carboxypeptidase C
588	25	56.8	334	2	A10795	probable receptor/	661	25	56.8	501	2	E70596	hypothetical prote
589	25	56.8	335	2	A40036	apoptosis-mediati	662	25	56.8	503	2	D70930	hypothetical prote
590	25	56.8	339	2	E83209	conserved hypothet	663	25	56.8	509	2	T31136	multidrug-efflux t
591	25	56.8	341	2	G97458	hypothetical prote	664	25	56.8	509	2	A86578	hypothetical prote
592	25	56.8	342	2	F70311	conserved hypothet	665	25	56.8	510	1	S43516	carboxypeptidase C
593	25	56.8	345	2	A05279	collagen - Caenoh	666	25	56.8	516	2	G86454	CDS protein FYJ11.
594	25	56.8	348	2	A34705	hypothetical secre	667	25	56.8	526	2	AB3433	C-di-GMP phosphodi
595	25	56.8	363	2	E97223	hypothetical prote	668	25	56.8	527	2	T41856	AcNMPV orf119 - Bo
596	25	56.8	364	2	H71440	hypothetical prote	669	25	56.8	529	2	T10388	hypothetical prote
597	25	56.8	365	2	T06693	hypothetical prote	670	25	56.8	530	2	H72864	AcOrf-119 protein
598	25	56.8	365	2	T24955	hypothetical prote	671	25	56.8	531	2	S20900	titin - mouse (fra
599	25	56.8	370	2	C96979	similar to spore g	672	25	56.8	531	2	B84442	hypothetical prote
600	25	56.8	375	2	JX0131	cellulase (EC 3.2.	673	25	56.8	532	2	C96514	hypothetical prote
601	25	56.8	385	2	T22822	hypothetical prote	674	25	56.8	544	2	T16374	hypothetical prote
602	25	56.8	388	2	AB3913	xyylanase Y BH2105	675	25	56.8	547	2	T41613	membrane transport
603	25	56.8	394	2	S58301	hypothetical prote	676	25	56.8	548	2	T23137	hypothetical prote
604	25	56.8	395	2	T00516	hypothetical prote	677	25	56.8	550	2	T26562	hypothetical prote
605	25	56.8	397	2	S74344	citrate synthase g	678	25	56.8	553	2	T14354	probable somatic e
606	25	56.8	400	2	T46383	hypothetical prote	679	25	56.8	555	2	A82208	hypothetical prote
607	25	56.8	403	2	T06762	hypothetical prote	680	25	56.8	557	2	AD0249	probable hemolysin
608	25	56.8	404	2	S64480	hypothetical prote	681	25	56.8	578	2	A70877	steryl-sulfatase (
609	25	56.8	407	2	I48739	Mha2 (keratin acid	682	25	56.8	583	1	KJHUC	protein C05B5.5
610	25	56.8	407	2	A84538	hypothetical prote	683	25	56.8	584	2	JC7809	protein C05B5.5 [1
611	25	56.8	415	2	T49840	hypothetical prote	684	25	56.8	585	2	S43572	protein C05B5.5 [1
612	25	56.8	416	1	JN0006	nerve growth facto	685	25	56.8	585	2	B88571	protein kinase C (
613	25	56.8	418	1	S28372	cellulase (EC 3.2.	686	25	56.8	586	2	A53758	protein kinase C (

687	25	56.8	587	2	A49509	protein kinase C (760	25	56.8	6805	2	S20901	cltin - rabbit (fr
688	25	56.8	591	2	S50972	Ribz protein - yea	761	25	56.8	26926	1	I38344	cltin, cardiac mus
689	25	56.8	592	1	A30314	protein kinase C (762	24.5	56.7	72	1	CHBC4	heat-stable entero
690	25	56.8	592	1	UC1480	protein kinase C (763	24.5	56.7	72	1	CHBC1B	heat-stable entero
691	25	56.8	592	2	JN0877	protein kinase C (764	24.5	55.7	293	3	S31763	chitinase (EC 3.2.
692	25	56.8	594	2	G84441	probable sucrose-p	765	24.5	55.7	301	4	B40201	artifact-warning s
693	25	56.8	645	2	T05251	probable disease r	766	24.5	55.7	330	2	T36695	probable integral
694	25	56.8	647	2	T23814	hypothetical prote	767	24.5	55.7	600	2	D84258	alcohol dehydrogen
695	25	56.8	651	2	UC7705	death receptor-6 -	768	24.5	55.7	802	2	T24293	hypothetical prote
696	25	56.8	668	2	T05257	probable disease r	769	24.5	55.7	901	2	T04108	receptor kinase ho
697	25	56.8	670	2	F88297	protein M28.1 [imp	770	24.5	55.7	949	2	T24294	hypothetical prote
698	25	56.8	679	2	AD0507	probable glycosyl	771	24.5	55.7	1042	2	A57534	mucin 5AC, (clone L
699	25	56.8	704	1	AD0507	DNA-directed DNA p	772	24.5	55.7	1056	2	AS3767	pseudo-hevein - pa
700	25	56.8	704	2	S07512	DNA-directed DNA p	773	24	54.5	45	2	S50180	metallothionein pr
701	25	56.8	709	2	S33907	glycerol kinase (E	774	24	54.5	61	2	S14049	metallothionein-1i
702	25	56.8	717	2	E86812	sugar hydrolase (i	775	24	54.5	66	2	T17015	metallothionein-20
703	25	56.8	732	2	B69749	hypothetical prote	776	24	54.5	71	2	S47576	metallothionein 20
704	25	56.8	741	2	T05250	probable disease r	777	24	54.5	71	2	S39422	metallothionein 20
705	25	56.8	769	2	A71403	probable transport	778	24	54.5	71	2	S39421	metallothionein 20
706	25	56.8	771	2	AF2381	hypothetical prote	779	24	54.5	71	2	S39420	metallothionein 20
707	25	56.8	781	2	T20864	hypothetical prote	780	24	54.5	71	2	S47577	metallothionein 20
708	25	56.8	835	2	T05259	probable disease r	781	24	54.5	72	2	S39418	metallothionein 10
709	25	56.8	844	2	B64678	NADH2 dehydrogenas	782	24	54.5	72	2	S39419	metallothionein 10
710	25	56.8	846	2	H70599	hypothetical prote	783	24	54.5	72	2	S39417	metallothionein 10
711	25	56.8	849	2	H71838	NADH2 dehydrogenas	784	24	54.5	72	2	S39416	metallothionein 10
712	25	56.8	852	2	C98310	haer protein (Y089	785	24	54.5	81	2	S27366	metallothionein B (
713	25	56.8	868	2	D86349	hypothetical prote	786	24	54.5	81	2	S27369	metallothionein B (
714	25	56.8	869	2	A71400	probable disease r	787	24	54.5	95	2	S76543	hypothetical prote
715	25	56.8	910	2	T29935	hypothetical prote	788	24	54.5	102	1	KREUB	keratin, feather -
716	25	56.8	919	2	T32541	unc-5 protein - Ca	789	24	54.5	115	2	S16534	metallothionein-1i
717	25	56.8	923	2	AH2972	heme receptor haer	790	24	54.5	120	2	F82185	conserved hypotet
718	25	56.8	947	1	B44294	unc-5 protein, lon	791	24	54.5	122	2	JC6548	high sulfur protei
719	25	56.8	947	1	T43236	carboxypeptidase C	792	24	54.5	134	2	T22275	hypothetical prote
720	25	56.8	1002	2	T43236	genome polypeptid	793	24	54.5	168	2	S78110	chi protein - Rhiz
721	25	56.8	1018	1	GNMXG7	laminin B1k chain	794	24	54.5	174	2	B90657	hypothetical lipop
722	25	56.8	1170	2	AS3612	laminin gamma 2 ch	795	24	54.5	174	2	B85508	hypothetical prote
723	25	56.8	1192	2	S69000	laminin B2t chain	796	24	54.5	177	2	H97205	hypothetical prote
724	25	56.8	1231	2	T35227	probable nitrate r	797	24	54.5	183	2	T49855	hypothetical prote
725	25	56.8	1245	2	T42820	hypothetical prote	798	24	54.5	195	2	T28803	hypothetical prote
726	25	56.8	1246	2	C36806	hypothetical prote	799	24	54.5	200	2	F69047	hypothetical prote
727	25	56.8	1302	2	T00038	probable RNA-direc	800	24	54.5	201	2	D71190	hypothetical prote
728	25	56.8	1309	2	T00078	hypothetical prote	801	24	54.5	202	2	T32159	hypothetical prote
729	25	56.8	1321	2	T00382	hypothetical prote	802	24	54.5	235	2	D65071	hypothetical prote
730	25	56.8	1376	2	G00043	osteochondogen - hum	803	24	54.5	235	2	A85943	hypothetical prote
731	25	56.8	1378	2	T47505	RING finger-like p	804	24	54.5	235	2	E91097	hypothetical prote
732	25	56.8	1490	2	UC5145	DNA (cytosine-5-) -	805	24	54.5	241	1	CCBO1	ubiquinol-cytochro
733	25	56.8	1495	2	S22610	DNA (cytosine-5-) -	806	24	54.5	245	2	F90823	probable tail asse
734	25	56.8	1502	1	RGBYH1	CYC1/CYP3 transcri	807	24	54.5	245	2	C85682	probable tail asse
735	25	56.8	1537	2	JC4172	DNA (cytosine-5-) -	808	24	54.5	247	2	B90877	probable tail asse
736	25	56.8	1550	2	S60228	glutamate synthase	809	24	54.5	247	2	B90997	tail assembly prot
737	25	56.8	1607	1	MMMSB2	laminin gamma-1 ch	810	24	54.5	247	2	F90834	hypothetical prote
738	25	56.8	1609	1	MMHTB2	laminin gamma-1 ch	811	24	54.5	270	2	T47421	hypothetical prote
739	25	56.8	1612	2	UC5210	DNA (cytosine-5-) -	812	24	54.5	272	2	H83269	hypothetical prote
740	25	56.8	1622	2	JEC0378	adhesive ligand ep	813	24	54.5	272	2	T24709	hypothetical prote
741	25	56.8	1713	2	A55347	MEGF3 protein - hu	814	24	54.5	275	2	B55224	hypothetical prote
742	25	56.8	1737	2	T00209	laminin alpha-2 ch	815	24	54.5	280	2	T09504	LIM-protein FHL3,
743	25	56.8	1751	1	MMHTMH	hypothetical prote	816	24	54.5	282	1	YPDOD1	prestalk D11 prote
744	25	56.8	1808	2	T15099	hypothetical prote	817	24	54.5	289	2	F71402	xyloglucan endo-1,
745	25	56.8	1945	2	T13337	Plexin A - fruit f	818	24	54.5	289	2	S04648	ampr protein - Rho
746	25	56.8	2116	2	T49818	glutamate synthase	819	24	54.5	289	2	T49024	hypothetical prote
747	25	56.8	2150	2	T32497	hypothetical prote	820	24	54.5	293	2	A85817	hypothetical prote
748	25	56.8	2356	2	T27790	hypothetical prote	821	24	54.5	293	2	E85718	probable tail comp
749	25	56.8	2355	2	T02546	hypothetical prote	822	24	54.5	300	2	T49748	hypothetical prote
750	25	56.8	2823	2	F87908	protein T22A3.8 [i	823	24	54.5	302	2	AB0857	ATP sulfurylase (A
751	25	56.8	2823	2	T23064	hypothetical prote	824	24	54.5	302	2	D65056	ATP sulfurylase su
752	25	56.8	3075	2	S14458	laminin alpha-1 ch	825	24	54.5	302	2	G85924	ATP sulfurylase su
753	25	56.8	3084	1	MMMSA	laminin alpha-1 ch	826	24	54.5	302	2	F91079	sulfate adenylylitr
754	25	56.8	3102	2	T43291	laminin alpha-1 ch	827	24	54.5	302	2	AH0408	hypothetical prote
755	25	56.8	3106	1	S53868	laminin alpha-2 ch	828	24	54.5	302	2	T32872	hypothetical prote
756	25	56.8	3672	2	T23433	hypothetical prote	829	24	54.5	303	2	A86443	probable major int
757	25	56.8	3704	2	T37316	probable laminin a	830	24	54.5	313	2	A11948	hypothetical prote
758	25	56.8	3707	2	S18252	heparan sulfate pr	831	24	54.5	320	2	A53119	cell adhesion glyco
759	25	56.8	4391	2	A38096	perlecan precursor	832	24	54.5	325	1	S00680	ubiquinol-cytochro

833	24	54.5	34.1	2	T04454	probable peroxidase	906	23.5	53.4	492	2	A61382	phosphorylation re
834	24	54.5	34.7	2	T41962	hypothetical prote	907	23.5	53.4	527	2	C64139	translation releas
835	24	54.5	355	2	S74204	C4-dicarboxylate m	908	23.5	53.4	529	2	AD1072	peptide chain rele
836	24	54.5	356	2	T04360	probable peroxidase	909	23.5	53.4	529	2	E91295	peptide chain rele
837	24	54.5	353	2	C64098	crRNA (uracil-5')-m	910	23.5	53.4	529	2	G86136	peptide chain rele
838	24	54.5	366	2	A37321	crRNA (uracil-5')-m	911	23.5	53.4	529	2	AE0053	peptide chain rele
839	24	54.5	366	2	H91240	crRNA (uracil-5')-m	912	23.5	53.4	529	2	I59305	translation releas
840	24	54.5	366	2	AC0935	crRNA (uracil-5')-me	913	23.5	53.4	531	2	E82295	translation releas
841	24	54.5	366	2	EB6088	crRNA (uracil-5')-m	914	23	52.3	27	2	T12330	metallochionein -
842	24	54.5	367	2	AE0476	crRNA (uracil-5')-m	915	23	52.3	46	1	VTVA43	viscotoxin A3 - Bu
843	24	54.5	373	2	T29596	hypothetical prote	916	23	52.3	46	1	DKOGB	viscotoxin A3 - Bu
844	24	54.5	380	2	D44490	retrovirus-related	917	23	52.3	46	1	LORDAL	ligetoxin A - Arge
845	24	54.5	384	2	D69809	spore germination	918	23	52.3	46	1	VFPDIT	phoratoxin A - Arge
846	24	54.5	392	2	B48423	homeotic protein e	919	23	52.3	46	1	VTVA42	viscotoxin A2 - Bu
847	24	54.5	401	2	A48423	engrailed homeodom	920	23	52.3	46	1	VTVAAB	viscotoxin B - Eur
848	24	54.5	407	2	T41708	gtp binding protei	921	23	52.3	52	2	D95920	hypothetical prote
849	24	54.5	427	2	S71367	u1 snRNP 70K prote	922	23	52.3	57	1	TIRPVB	venom basic protei
850	24	54.5	435	2	T08038	probable mitochond	923	23	52.3	59	2	I51272	retinoic acid rece
851	24	54.5	450	2	G01158	tyrosine kinase ac	924	23	52.3	63	2	S07127	chymotrypsin/elast
852	24	54.5	468	2	T03164	probable major env	925	23	52.3	65	2	T09787	metallochionein-11
853	24	54.5	484	2	AC3022	glutamate synthase	926	23	52.3	65	4	C40499	defensin alpha-3 p
854	24	54.5	484	2	G98262	glutamate synthase	927	23	52.3	67	1	TIROC	trypsin inhibitor
855	24	54.5	499	1	P1MRJ3	li protein - human	928	23	52.3	75	1	GSRF8	salivary glue prot
856	24	54.5	499	2	AE3514	glutamate synthase	929	23	52.3	80	2	AH2958	hypothetical prote
857	24	54.5	506	1	CSFP	catalase (EC 1.11.	930	23	52.3	80	2	F96324	hypothetical prote
858	24	54.5	507	2	S64507	probable membrane	931	23	52.3	97	2	T38323	very hypothetical
859	24	54.5	516	2	T35542	probable large sec	932	23	52.3	97	2	B25439	Beein protein - s
860	24	54.5	521	2	H69317	conserved hypotet	933	23	52.3	105	2	F72614	hypothetical prote
861	24	54.5	524	1	P1MR58	li protein - human	934	23	52.3	111	2	S16099	viscotoxin - Europ
862	24	54.5	532	1	C97228	probable peptide A	935	23	52.3	111	2	E72648	hypothetical prote
863	24	54.5	538	2	G69317	conserved hypotet	936	23	52.3	122	2	U00150	hypothetical 13K p
864	24	54.5	565	2	T14732	probable beta-gluc	937	23	52.3	133	2	S52547	thionin variant Th
865	24	54.5	602	1	S25116	hydroxymethylgluta	938	23	52.3	134	2	S52554	thionin variant Th
866	24	54.5	618	1	B43255	hydrogenase (EC 1.	939	23	52.3	134	2	S52553	thionin variant Th
867	24	54.5	618	2	S09251	hydrogenase (EC 1.	940	23	52.3	134	2	T49623	hypothetical prote
868	24	54.5	619	1	UH0776	hydrogenase (EC 1.	941	23	52.3	135	2	H84523	probable thionin l
869	24	54.5	622	1	OYCHGC	phosphoenolpyruvat	942	23	52.3	136	2	I56339	cytochrome-c oxida
870	24	54.5	622	1	OYRTGP	phosphoenolpyruvat	943	23	52.3	141	2	S73027	hypothetical prote
871	24	54.5	622	1	A45746	phosphoenolpyruvat	944	23	52.3	148	2	F71454	hypothetical prote
872	24	54.5	680	2	A28121	major merocrite su	945	23	52.3	152	1	KRSHHC	keratin high-sulfu
873	24	54.5	820	2	C81252	probable NMDH2 deh	946	23	52.3	152	2	I47111	high-sulfur wool m
874	24	54.5	829	2	A40894	RNA-directed RNA p	947	23	52.3	152	2	I47109	high-sulfur wool m
875	24	54.5	886	2	A57172	probable hormone r	948	23	52.3	152	2	I47108	high-sulfur wool m
876	24	54.5	912	2	UH0563	metabotropic gluta	949	23	52.3	152	2	I47112	high-sulfur wool m
877	24	54.5	956	1	A46016	thrombospondin 3 -	950	23	52.3	156	1	KRSHHB	keratin high-sulfu
878	24	54.5	956	2	A57121	thrombospondin 3 p	951	23	52.3	157	2	JN0057	hypothetical 17K p
879	24	54.5	957	2	T15976	hypothetical prote	952	23	52.3	158	2	T12958	copper transport p
880	24	54.5	1036	2	T38734	hypothetical prote	953	23	52.3	160	2	A69849	hypothetical prote
881	24	54.5	1043	2	T19734	hypothetical prote	954	23	52.3	162	2	I47107	high-sulfur wool m
882	24	54.5	1166	1	S06142	protein-tyrosine k	955	23	52.3	162	2	C81808	hypothetical prote
883	24	54.5	1173	1	VG1HHC	E2 glycoprotein pr	956	23	52.3	163	2	H97034	hypothetical secre
884	24	54.5	1225	1	S24284	E2 glycoprotein pr	957	23	52.3	165	2	B71432	hypothetical prote
885	24	54.5	1225	1	A36607	E2 glycoprotein -	958	23	52.3	169	1	S18946	ultra high-sulfur
886	24	54.5	1225	2	S55598	tegument protein 0	959	23	52.3	172	1	KRSHHA	keratin high-sulfu
887	24	54.5	1360	2	T33922	hypothetical prote	960	23	52.3	172	2	I47106	high-sulfur wool m
888	24	54.5	1414	2	T33236	hypothetical prote	961	23	52.3	179	2	T49508	hypothetical prote
889	24	54.5	1447	1	VG1HE3	E2 glycoprotein pr	962	23	52.3	182	1	KRSHHD	keratin high-sulfu
890	24	54.5	1447	1	VG1HE2	E2 glycoprotein pr	963	23	52.3	182	2	I47105	ultra-high-sulfur
891	24	54.5	1449	1	A43573	E2 glycoprotein pr	964	23	52.3	182	2	A36686	hypothetical prote
892	24	54.5	1449	1	VG1HFS	E2 glycoprotein pr	965	23	52.3	186	2	T32656	hypothetical prote
893	24	54.5	1449	2	S47423	E2 glycoprotein pr	966	23	52.3	187	2	T38452	hypothetical prote
894	24	54.5	1451	1	JOL1719	E2 glycoprotein pr	967	23	52.3	191	2	I46412	keratin KAP5.4 - s
895	24	54.5	1453	1	S41453	spike protein - ca	968	23	52.3	195	2	T03522	bluf protein - Rho
896	24	54.5	1506	2	T30886	integumentary muc	969	23	52.3	195	2	S52221	conserved hypotet
897	24	54.5	1534	2	T30295	p-glycoprotein - T	970	23	52.3	196	2	D69275	keratin KAP5.5 - s
898	24	54.5	1565	2	AD2135	polyketide synthas	971	23	52.3	197	2	I46413	T-cell surface gly
899	24	54.5	1587	2	G86467	hypothetical prote	972	23	52.3	198	2	S26566	hypothetical prote
900	24	54.5	1721	2	T21214	hypothetical prote	973	23	52.3	199	2	E75463	hypothetical prote
901	24	54.5	1772	2	A45532	major merocrite su	974	23	52.3	205	2	T14745	hypothetical prote
902	23	53.4	178	2	T32058	hypothetical prote	975	23	52.3	208	2	T34512	hypothetical prote
903	23	53.4	333	2	T32058	hypothetical prote	976	23	52.3	219	2	T21736	hypothetical prote
904	23	53.4	464	2	S50592	hypothetical prote	977	23	52.3	235	1	RWHUT8	T-cell surface gly
905	23	53.4	468	2	S39832	probable phosphopr	978	23	52.3	238	2	I48605	insulin-like growt

979	23	52.3	241	2	T29182
980	23	52.3	244	2	UC8019
981	23	52.3	247	1	D64740
982	23	52.3	250	1	DGBEL5
983	23	52.3	258	2	T44251
984	23	52.3	260	1	OTXL3
985	23	52.3	260	2	T11133
986	23	52.3	261	1	OTB03
987	23	52.3	261	1	OTC03
988	23	52.3	261	1	OTWS3
989	23	52.3	261	2	T11499
990	23	52.3	261	2	H34012
991	23	52.3	261	2	B30396
992	23	52.3	261	2	T30396
993	23	52.3	261	2	T11540
994	23	52.3	261	2	G58888
995	23	52.3	261	2	S47876
996	23	52.3	261	2	T11306
997	23	52.3	261	2	T11434
998	23	52.3	261	2	T11447
999	23	52.3	261	2	S60277
1000	23	52.3	261	2	D30396

ALIGNMENTS

RESULT 1
A27020
DIF-induced prestalk pdd63 protein precursor - slime mold (Dictyostelium discoideum) (F1
C/Species: Dictyostelium discoideum
C/Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004
C/Accession: A27020
R.Williams, J.G.; Ceccarelli, A.; McRobbie, S.; Mahbubani, H.; Kay, R.R.; Barly, A.; Ben
Cell 49, 185-192, 1987
A/Title: Direct induction of Dictyostelium prestalk gene expression by DIF provides evid
A/Reference number: A27020; PMID:87187613; PMID:3568124
A/Accession: A27020
A/Molecule type: DNA
A/Residues: 1-601 <WIL>
A/Cross-references: UNIPROT:Q7M4U3; UNIPARC:UPI000017B181
C/Genetics:
A/Genes: pdd63
F.1-26/Domains: signal sequence #status predicted <SIG>
F.21-601/Product: DIF-induced prestalk pdd63 protein #status predicted <MAT>

Query Match 81.8%; Score 36; DB 2; Length 601;
Best Local Similarity 71.4%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 245 CDPSTGC 251

RESULT 2
S07983
phospholipase A2 homolog precursor - common tiger snake
C/Species: Notechis scutatus scutatus (common tiger snake, mainland tiger snake)
C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 05-Oct-2004
C/Accession: S07983
R.Ducancel, F.
Submitted to the EMBL Data Library, January 1989
A/Reference number: S07983
A/Accession: S07983
A/Molecule type: mRNA
A/Residues: 1-145 <DUC>
A/Cross-references: UNIPROT:P20146; UNIPARC:UPI0000131165; EMBL:X14043; NID:964109; PIDD
C/Suprafamily: Phospholipase A2
F.1-27/Domains: signal sequence #status predicted <SIG>
F.28-145/Product: phospholipase A2 #status predicted <MAT>

Query Match 77.3%; Score 34; DB 2; Length 145;

Best Local Similarity 71.4%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 105 CDPSTGC 111

RESULT 3
T05348
hypothetical protein F8B4.80 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: T05348
R.Bevan, M.; Terry, N.; Ardles, W.; Buysaert, C.; Dasseville, R.; De Clerck, R.; De
ewes, H.W.; Mayer, K.F.X.; Schueller, C.
Submitted to the Protein Sequence Database, February 1999
A/Reference number: T05348
A/Accession: T05348
A/Molecule type: DNA
A/Residues: 1-312 <BEV>
A/Cross-references: UNIPROT:Q9SU73; UNIPARC:UPI0000048A40; EMBL:AL034567
A/Experimental source: cultivar Columbia; BAC clone F8B4
C/Genetics:
A/Map position: 4
A/Introns: 1/3; 44/3; 101/3; 139/3; 180/3
A/Note: F8B4.80

Query Match 77.3%; Score 34; DB 2; Length 312;
Best Local Similarity 71.4%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 208 CDPSTGC 214

RESULT 4
T50496
hypothetical protein T22D6.50 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C/Accession: T50496
R.Bevan, M.; Terry, N.; Ardles, W.; Buysaert, C.; Dasseville, R.; De Clerck, R.; De
ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
Submitted to the Protein Sequence Database, May 2000
A/Reference number: T22D6.50
A/Accession: T50496
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1058 <BEV>
A/Cross-references: UNIPROT:Q9LEZ5; UNIPARC:UPI000009F613; EMBL:AL357612
A/Experimental source: cultivar Columbia; BAC clone T22D6
C/Genetics:
A/Map position: 5
A/Introns: 60/3; 195/3; 222/3; 448/3; 492/3; 526/3; 555/2; 591/2; 616/3; 662/2; 715/3; 7
A/Note: T22D6.50

Query Match 75.0%; Score 33; DB 2; Length 1058;
Best Local Similarity 71.4%; Pred. No. 13e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 1019 CDPSTGC 1025

RESULT 5
F96006
hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) magaplaemid pSymB
C/Species: Sinorhizobium meliloti
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C/Accession: F96006

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernat
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:2136508; PMID:11481431
A:Accession: F96006
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-64 <KIR>
A:Cross-references: UNIPROT:Q92U25; UNIPARC:UPI00000CB812; GB:AL591985; P1DN:CAC49718.1
A:Experimental source: strain 1021, megaplasmid pSymB
R:Gallbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Weller, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Smb21688
A:Genome: plasmid

Query Match 72.7%; Score 32; DB 2; Length 64;
Best Local Similarity 57.1%; Pred. No. 22;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 4 CXPXTC 10
DB 15 CAPSGC 21

RESULT 6
T48397
S-Receptor kinase-like protein - Arabidopsis thaliana
N:Alternate names: protein F17C15.120
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 31-Dec-2004
C:Accession: T48397
R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24492
A:Accession: T48397
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-482 <BEV>
A:Cross-references: UNIPROT:Q9LZR8; UNIPARC:UPI000048ABC; EMBL:AL162506
A:Experimental source: cultivar Columbia; BAC clone F17C15
C:Genetics:
A:Map position: 5
A:Note: F17C15.120
C:Superfamily: S-locus glycoprotein

Query Match 70.5%; Score 31; DB 2; Length 482;
Best Local Similarity 57.1%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 4 CXPXTC 10
DB 301 CTFPGSC 307

RESULT 7
T25473
Hypothetical protein B0507.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25473
R:Bradsaw, H.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid B0507.
A:Reference number: Z20039
A:Accession: T25473

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-643 <BRA>
A:Cross-references: UNIPROT:Q22886; UNIPARC:UPI000017B6F9; EMBL:U64833; P1DN:AB04822.1
A:Experimental source: strain Bristol N2; clone B0507
C:Genetics:
A:Gene: CBSP:B0507.1
A:Map position: 5
A:Introns: 59/3; 133/1; 464/3; 586/1

Query Match 70.5%; Score 31; DB 2; Length 643;
Best Local Similarity 57.1%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 4 CXPXTC 10
DB 209 CSPQSGC 215

RESULT 8
A24255
Chorion class A protein L11 precursor - silkworm
C:Species: Bombyx mori (silkworm)
C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C:Accession: A24255
R:Spoerel, N.; Nguyen, H.T.; Kafatos, F.C.
J. Mol. Biol. 190, 23-35, 1986
A:Title: Gene regulation and evolution in the chorion locus of Bombyx mori. Structural
A:Reference number: A92929; MUID:87060979; PMID:3023635
A:Accession: A24255
A:Molecule type: DNA
A:Residues: 1-129 <SPO>
A:Cross-references: UNIPROT:P08826; UNIPARC:UPI0000127760; GB:X15557; GB:X04028; GB:X040
C:Superfamily: chorion class A protein pc292
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-129/Product: chorion class A protein L11 #status predicted <MAT>

Query Match 68.2%; Score 30; DB 2; Length 129;
Best Local Similarity 57.1%; Pred. No. 88;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 4 CXPXTC 10
DB 116 CAPTCGC 122

RESULT 9
B24255
Chorion class A protein L12 precursor - silkworm
C:Species: Bombyx mori (silkworm)
C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C:Accession: B24255
R:Spoerel, N.; Nguyen, H.T.; Kafatos, F.C.
J. Mol. Biol. 190, 23-35, 1986
A:Title: Gene regulation and evolution in the chorion locus of Bombyx mori. Structural
A:Reference number: A92929; MUID:87060979; PMID:3023635
A:Accession: B24255
A:Molecule type: DNA
A:Residues: 1-132 <SPO>
A:Cross-references: UNIPROT:P08825; UNIPARC:UPI0000127761; GB:X15557; GB:X04028; GB:X040
C:Superfamily: chorion class A protein pc292
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-132/Product: chorion class A protein L12 #status predicted <MAT>

Query Match 68.2%; Score 30; DB 2; Length 132;
Best Local Similarity 57.1%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 4 CXPXTC 10
DB 117 CAPTCGC 123

RESULT 10

AB0331
sigma E factor regulatory protein rsec [imported] - Yersinia pestis (strain C092)
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C/Accession: AB0331
R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; MUID:21470413; PMID:11586360
A/Accession: AB0331
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-151 <KUR>
A/Cross-references: UNIPROT:Q8ZD76; UNIPARC:UPI00000DC88B; GB:AL590842; PIDN:CMC92953.1;
C/Genetics:
A/Gene: rsec
C/Superfamily: Escherichia coli sigma-E factor regulatory protein rsec

Query Match 68.2%; Score 30; DB 2; Length 151;
Best Local Similarity 57.1%; Pred. No. 99;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
| | | |
DB 20 CEPHSGC 26

RESULT 11

T99486
hypothetical protein B14D6.380 [imported] - Neurospora crassa
C/Species: Neurospora crassa
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
C/Accession: T94986
R/Schulte, U.; Algen, V.; Hohenisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A/Reference number: Z25022
A/Accession: T94986
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-199 <SCH>
A/Cross-references: UNIPARC:UPI0000179E18; EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.380
A/Experimental source: BAC clone B14D6; strain OR74A
C/Genetics:
A/Gene: NCSP:B14D6.380
A/Map position: 6
C/Superfamily: Neurospora crassa hypothetical protein B14D6.380

Query Match 68.2%; Score 30; DB 2; Length 199;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
| | | |
DB 155 CKPLSGC 161

RESULT 12

S04926
wound-induced protein 1 precursor - potato
C/Species: Solanum tuberosum (potato)
C/Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C/Accession: S04926
R/Stanford, A.; Bevan, M.; Northcote, D.
Mol. Gen. Genet. 215, 200-208, 1989
A/Title: Differential expression within a family of novel wound-induced genes in potato.
A/Reference number: S04926; MUID:89218921; PMID:2710059
A/Molecule type: DNA
A/Residues: 1-200 <STA>
A/Cross-references: UNIPROT:P09761; UNIPARC:UPI0000138F0C; EMBL:X13497; NID:G21617; PIDN

C/Genetics:

A/Gene: win1
A/Accession: 143/3
C/Superfamily: hevein precursor; barwin homology; hevein chitin-binding domain homology
F/1-25/Domain: signal sequence #status predicted <SIG>
F/26-200/Product: wound-induced protein 1 #status predicted <MAT>
F/26-69/Domain: hevein chitin-binding domain homology <HCB>
F/78-199/Domain: barwin homology <BAR>

Query Match 68.2%; Score 30; DB 2; Length 200;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
| | | |
DB 56 CSPSGC 62

RESULT 13

T07729
wound-induced protein (clone TAB7) - tomato (fragment)
C/Species: Lycopersicon esculentum (tomato)
C/Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C/Accession: T07729
R/Harris, N.; Taylor, J.E.; Roberts, J.A.
J. Exp. Bot. 48, 1223-1227, 1997
A/Title: Characterization and expression of an mRNA encoding a wound-induced (win) prote
A/Reference number: Z16099
A/Accession: T07729
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-201 <HAR>
A/Cross-references: UNIPROT:003994; UNIPARC:UPI00000AC9P2; EMBL:U89764; NID:G1888560; PI
A/Experimental source: strain Alisa Craig; leaf abscission zone tissue
C/Superfamily: hevein precursor; barwin homology; hevein chitin-binding domain homology
F/16-60/Domain: hevein chitin-binding domain homology <HCB>
F/68-188/Domain: barwin homology <BAR>

Query Match 68.2%; Score 30; DB 2; Length 201;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
| | | |
DB 46 CSPSGC 52

RESULT 14

T04098
CBP20 preproprotein - common tobacco
N/Alternate names: wound-induced protein
C/Species: Nicotiana tabacum (common tobacco)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: T04098
R/Ponstein, A.S.; Bres-Vloemans, S.A.; Sela-Buurlage, W.B.; Elzen, P.J.; Meijchers, L.S.,
Plant Physiol. 104, 109-118, 1994
A/Title: A novel pathogen- and wound-inducible tobacco (Nicotiana tabacum) protein with
A/Reference number: Z15209; MUID:94159785; PMID:8115541
A/Accession: T04098
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-211 <PON>
A/Cross-references: UNIPROT:Q41231; UNIPARC:UPI00000A8993; EMBL:S72452; NID:G632733; PID
A/Experimental source: cultivar Samum NN
C/Superfamily: hevein precursor; barwin homology; hevein chitin-binding domain homology
F/23-66/Domain: hevein chitin-binding domain homology <HCB>
F/78-199/Domain: barwin homology <BAR>

Query Match 68.2%; Score 30; DB 2; Length 211;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10

Db 53 CXPXGTC 59

RESULT 15

S04927
wound-induced protein 2 precursor - potato
C/Species: Solanum tuberosum (potato)
C/Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C/Accession: S04927
R/Stanford, A.; Bevan, M.; Northcote, D.
Mol. Gen. Genet. 215, 200-208, 1989
A/Title: Differential expression within a family of novel wound-induced genes in potato.
A/Reference number: S04926; MUID:89218921; PMID:2710099
A/Accession: S04927
A/Molecule type: DNA
A/Residues: 1-211 <STA>
A/Cross-references: UNIPROT:P09762; UNIPARC:UPI0000138F0D; EMBL:X13497; NID:g21617; PIDN
C/Genetics:
A/Genes: win2
A/Intons: 142/3
C/Suprafamily: hevein precursor; barwin homology; hevein chitin-binding domain homology
F/1-25/Domain: signal sequence #status predicted <SIG>
F/26-211/Product: wound-induced protein 2 #status predicted <MAT>
F/26-69/Domain: hevein chitin-binding domain homology <HCB>
F/77-198/Domain: barwin homology <BAR>

Query Match 68.2%; Score 30; DB 2; Length 211;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXGTC 10
Db 56 CXPXGTC 62

RESULT 16

MMVZB4
anthrombin-III homolog 3 - vaccinia virus (strain Copenhagen)
N/Alternate names: B14R protein; serine proteinase inhibitor
C/Species: vaccinia virus
A/Note: host Homo sapiens (man)
C/Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 05-Oct-2004
C/Accession: D42527
R/Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Davis, S.W.; Winslow, J.P.; Paolletti, E.
Virology 179, 517-563, 1990
A/Title: Appendix to "The complete DNA sequence of vaccinia virus".
A/Reference number: A42501
A/Accession: D42527
A/Molecule type: DNA
A/Residues: 1-222 <GOE>
A/Cross-references: UNIPROT:P20842; UNIPARC:UPI0000135E37; GB:M35027; NID:g335317; PIDD:
A/Experimental source: strain Copenhagen
R/Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Davis, S.W.; Winslow, J.P.; Paolletti, E.
Virology 179, 247-266, 1990
A/Title: The complete DNA sequence of vaccinia virus.
A/Reference number: A42531; MUID:91021027; PMID:2219722
A/Contents: annotation; possible protein-coding frames
A/Note: neither amino acid nor nucleotide sequence is given
C/Suprafamily: serpin
C/Keywords: glycoprotein; serine proteinase inhibitor
F/62,93/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 68.2%; Score 30; DB 1; Length 222;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXGTC 10
Db 216 CXPXGTC 222

RESULT 17

E84326
hypothetical protein Vmg1748c [imported] - Halobacterium sp. NRC-1

C/Species: Halobacterium sp. NRC-1
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 05-Oct-2004
C/Accession: E84326
R/Ng, W.V.; Kennedy, S.P.; Mahatae, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laeky, S.;
Leithauer, B.; Keller, K.; Cruz, R.; Danison, M.J.; Hough, D.W.; Maddocks, D.G.; Jabic
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
A/Title: Genome sequence of Halobacterium species NRC-1.
A/Reference number: E84326
A/Accession: E84326
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-224 <STO>
A/Cross-references: UNIPROT:Q9HP93; UNIPARC:UPI0000063970; GB:AE004437; NID:g10581204; I
C/Genetics:
A/Genes: VMG1748C
C/Suprafamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 68.2%; Score 30; DB 2; Length 224;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXGTC 10
Db 168 CXPXGTC 174

RESULT 18

A83801
ABC transporter (substrate-binding protein) BH1209 [imported] - Bacillus halodurans (str
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: A83801
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: A83801
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-336 <STO>
A/Cross-references: UNIPROT:Q9KDK4; UNIPARC:UPI00000D73E9; GB:AP001511; GB:BA000004; NID
A/Experimental source: strain C-125
C/Genetics:
A/Genes: BH1209

Query Match 68.2%; Score 30; DB 2; Length 336;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXGTC 10
Db 139 CXPXGTC 145

RESULT 19

MMVZW2
anthrombin-III homolog 2 - vaccinia virus
N/Alternate names: B13R protein; serine proteinase inhibitor 2
C/Species: vaccinia virus
C/Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 05-Oct-2004
C/Accession: A44035; J01807; B30175; C42527
R/Smith, G.D.; Howard, S.T.; Chan, Y.S.
J. Gen. Virol. 70, 2333-2343, 1989
A/Title: Vaccinia virus encodes a family of genes with homology to serine proteinase inh
A/Reference number: A34035; MUID:89381686; PMID:2778436
A/Accession: A34035
A/Molecule type: DNA
A/Residues: 1-345 <SMT>
A/Cross-references: UNIPROT:P15059; UNIPARC:UPI0000135B2A; EMBL:D00581; NID:g222694; PID

RESULT 19

A:Experimental source: strain WR
 R:Smith, G.L.; Chan, Y.S.; Howard, S.T.
 J:Gen. Virol. 72, 1349-1376, 1991
 A:Title: Nucleotide sequence of 42kbp of vaccinia virus strain WR from near the right in
 A:Reference number: JQ1767; MUID:91259063; PMID:2045793
 A:Accession: JQ1807
 A:Molecule type: DNA
 A:Residues: 1-345 <SN>
 A:Cross-references: UNIPARC:UPI0000135E2A; DDBJ:D11079; NID:g222717; PIDN:BAA01843.1; PI
 A:Experimental source: strain WR
 R:Kotwal, G.J.; Moss, B.
 J:Virol. 63, 600-606, 1989
 A:Title: Vaccinia virus encodes two proteins that are structurally related to members of
 A:Reference number: A30175; MUID:89094985; PMID:2783466
 A:Accession: B30175
 A:Molecule type: DNA
 A:Residues: 1-56, 'V', 61-129, 'S', 131-133, 'DV', 136-164, 'CI', 167-171, 'E', 173-179, 'T', 181-19
 A:Cross-references: UNIPARC:UPI0000173205
 R:Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Davis, S.W.; Winslow, J.P.; Paolletti, E.
 Virol. 179, 517-563, 1990
 A:Title: Appendix to "The complete DNA sequence of vaccinia virus".
 A:Reference number: A42501
 A:Accession: C42527
 A:Molecule type: DNA
 A:Residues: 11-47, 'T', 49-51, 'T', 53-78, 'F', 80-81, 'G', 83-124, 'SN' <GOE>
 A:Cross-references: UNIPARC:UPI0000135E35; GB:M35027; NID:g335317; PIDN:AAA48210.1; PID:
 A:Experimental source: strain Copenhagen
 R:Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Davis, S.W.; Winslow, J.P.; Paolletti, E.
 Virol. 179, 247-266, 1990
 A:Title: The complete DNA sequence of vaccinia virus.
 A:Reference number: A42531; MUID:91021027; PMID:2219722
 A:Contents: annotation; possible protein-coding frames
 A:Note: neither amino acid nor nucleotide sequence is given
 C:Genetics:
 A:Gene: SPI-2
 C:Superfamily: serpin
 C:Keywords: glycoprotein; serine proteinase inhibitor
 F:35,59,185,216/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 68.2%; Score 30; DB 1; Length 345;
 Best Local Similarity 57.1%; Pred. No. 1.9e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTC 10
 DB 339 CSPXTC 345

RESULT 20
 T23898
 hypothetical protein R04D3.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T23898
 R:Swindburne, J.
 submitted to the EMBL Data Library, March 1996
 A:Reference number: Z19815
 A:Accession: T23898
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-406 <WIL>
 A:Cross-references: UNIPROT:Q21712; UNIPARC:UPI00000766B6; EMBL:Z70212; PIDN:CAA94162.1;
 A:Experimental source: clone R04D3
 C:Genetics:
 A:Gene: CSP:R04D3.2
 A:Map position: X
 A:Introns: 108/3; 320/3

Query Match 68.2%; Score 30; DB 2; Length 406;
 Best Local Similarity 57.1%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTC 10

DB 36 CSPNAGC 42

RESULT 21
 T23899
 hypothetical protein R04D3.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T23899
 R:Swindburne, J.
 submitted to the EMBL Data Library, March 1996
 A:Reference number: Z19815
 A:Accession: T23899
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-410 <WIL>
 A:Cross-references: UNIPROT:Q21713; UNIPARC:UPI000007D6C1; EMBL:Z70212; PIDN:CAA94163.1;
 A:Experimental source: clone R04D3
 C:Genetics:
 A:Gene: CSP:R04D3.3
 A:Map position: X
 A:Introns: 108/3; 296/3; 352/3

Query Match 68.2%; Score 30; DB 2; Length 430;
 Best Local Similarity 57.1%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTC 10
 DB 36 CSPNAGC 42

RESULT 22
 S01793
 arylsulphatase (EC 3.1.6.1) precursor - sea urchin (Hemicentrotus pulcherrimus)
 C:Species: Hemicentrotus pulcherrimus
 C>Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 05-Oct-2004
 C:Accession: S01793
 R:Sasaki, H.; Yamada, K.; Akasaka, K.; Kawasaki, H.; Suzuki, K.; Saito, A.; Sato, M.; Sh
 Eur. J. Biochem. 177, 9-13, 1988
 A:Title: cDNA cloning, nucleotide sequence and expression of the gene for arylsulphatase
 A:Reference number: S01793; MUID:89030699; PMID:3181160
 A:Accession: S01793
 A:Molecule type: mRNA
 A:Residues: 1-551 <SN>
 A:Cross-references: UNIPROT:P14000; UNIPARC:UPI00001260B5; EMBL:X17015; NID:g9432; PID:g
 A:Note: part of this sequence was confirmed by protein sequencing
 A:Note: the authors translated the codon CAC for residue 61 as Asp
 C:Superfamily: Animal sulfatase
 C:Keywords: sulfuric ester hydrolase
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:1-551/Product: arylsulphatase #status predicted <MT>
 F:100/Modified site: 3-oxoalanine (Cys) #status predicted

Query Match 68.2%; Score 30; DB 2; Length 551;
 Best Local Similarity 57.1%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTC 10
 DB 518 CNPANGC 524

RESULT 23
 S07089
 arylsulphatase (EC 3.1.6.1) - sea urchin (Lytechinus pictus)
 C:Species: Lytechinus pictus (painted urchin)
 C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 05-Oct-2004
 C:Accession: S07089
 R:Yamada, K.; Akasaka, K.; Shimada, H.
 Eur. J. Biochem. 186, 405-410, 1989
 A:Title: structure of sea-urchin arylsulphatase gene.

A:Reference number: S07089; MUID:90092130; PMID:2598936
A:Accession: S07089
A:Molecule type: DNA
A:Residues: 1-551 <YAM>
A:Cross-references: UNIPROT:P14000; UNIPARC:UPI00001759C4; EMBL:X16679
C:Genetics:
A:Gene: Ars
A:Introns: 103/2; 160/1; 197/3; 272/1; 327/1
C:Superfamily: Animal sulfatase
C:Keywords: sulfuric ester hydrolase
F:100/Modified site: 3-oxoalanine (Cys) #status predicted

Query Match 68.2%; Score 30; DB 2; Length 551;
Best Local Similarity 57.1%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 518 CNEPANGC 524

RESULT 24
T23864
hypothetical protein F39D8.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T23864
R:Henry, C.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19809
A:Accession: T23864
A:Status: preliminary; translated from GB/EMBL/DDBB
A:Molecule type: DNA
A:Residues: 1-646 <MTL>
A:Cross-references: UNIPROT:Q20191; UNIPARC:UPI000017B99F; EMBL:Z69793; PIDN:CAA93676.1;
C:Genetics:
A:Experimental source: clone R03A10
A:Gene: CESP:F39D8.4
A:Map position: X
A:Introns: 22/2; 112/3; 160/1; 250/1; 334/1; 424/3; 463/1; 485/3; 528/2; 593/2

Query Match 68.2%; Score 30; DB 2; Length 648;
Best Local Similarity 57.1%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 50 CLPTAGC 56

RESULT 25
E89753
Protein F11C7.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: E89753
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: E89753
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1172 <STO>
A:Cross-references: UNIPROT:Q19350; UNIPARC:UPI0000074975; GB:chr_X; PIDN:AAC69012.1; PI
C:Genetics:
A:Gene: F11C7.4
A:Map position: X

Query Match 68.2%; Score 30; DB 2; Length 1722;
Best Local Similarity 57.1%; Pred. No. 6.5e+02;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 401 CEPKSGC 407

RESULT 26
H71262
hypothetical protein TP0940 - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: H71262
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwiltson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Ullrich, T.; McPherson, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876
A:Accession: H71262
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-46 <COL>
A:Cross-references: UNIPROT:O83910; UNIPARC:UPI00001398A5; GB:AE001262; GB:AE000520; NIT
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0940

Query Match 65.9%; Score 29; DB 2; Length 46;
Best Local Similarity 57.1%; Pred. No. 60;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 36 CEPAPGC 42

RESULT 27
B45937
early chorion protein Era.1 precursor - silkworm
C:Species: Bombyx mori (silkworm)
C>Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 09-Jul-2004
C:Accession: B45937
R:Ritner, B.L.; Burke, W.D.; Lecanidou, R.; Rodakis, G.C.; Eickbush, T.H.
Dev. Biol. 125, 423-431, 1988
A:Title: Organization and expression of three genes from the silkworm early chorion loci
A:Reference number: A45937; MUID:88112521; PMID:3338621
A:Accession: B45937
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-119 <HIB>
A:Cross-references: UNIPROT:P13511; UNIPARC:UPI0000127783; GB:M19076; NID:G155974; PIDN
C:Superfamily: chorion class A protein pc293

Query Match 65.9%; Score 29; DB 2; Length 119;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 107 CGPTGCG 113

RESULT 28
S24292
chorion protein - silkworm
C:Species: Bombyx mori (silkworm)
C>Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C:Accession: S24292
R:Hibner, B.L.; Burke, W.D.; Eickbush, T.H.
Genetics 128, 595-606, 1991
A:Title: Sequence identity in an early chorion multigene family is the result of local
A:Reference number: S24291; MUID:91340109; PMID:1874417

A:Accession: S24292
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-119 <HIB>
A:Cross-references: UNIPROT:Q17213; UNIPARC:UPI0000127787; EMBL:X58446; NID:G5852; PIDN:
C:Genetics:
A:introns: 17/3
C:Superfamily: chorion class A protein pc292

Query Match 65.9%; Score 29; DB 2; Length 119;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 107 CGPTCGC 113

RESULT 29
S24294
chorion protein - silkworm
C:Species: Bombyx mori (silkworm)
C>Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C:Accession: S24294
R:Hitner, B.L.; Burke, W.D.; Eickbush, T.H.
Genetics 128, 595-606, 1991
A>Title: Sequence identity in an early chorion multigene family is the result of localiz
A:Reference number: S24291; MUID:91340109; PMID:1874417
A:Accession: S24294
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-119 <HIB>
A:Cross-references: UNIPROT:Q17214; UNIPARC:UPI0000127789; EMBL:X58448; NID:G5856; PIDN:
C:Genetics:
A:introns: 17/3
C:Superfamily: chorion class A protein pc292

Query Match 65.9%; Score 29; DB 2; Length 119;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 107 CGPTCGC 113

RESULT 30
S24291
chorion protein - silkworm
C:Species: Bombyx mori (silkworm)
C>Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C:Accession: S24291
R:Hitner, B.L.; Burke, W.D.; Eickbush, T.H.
Genetics 128, 595-606, 1991
A>Title: Sequence identity in an early chorion multigene family is the result of localiz
A:Reference number: S24291; MUID:91340109; PMID:1874417
A:Accession: S24291
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-119 <HIB>
A:Cross-references: UNIPROT:Q17212; UNIPARC:UPI0000127785; EMBL:X58445; NID:G5850; PIDN:
C:Genetics:
A:introns: 17/3
C:Superfamily: chorion class A protein pc292

Query Match 65.9%; Score 29; DB 2; Length 119;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 107 CGPTCGC 113

RESULT 31
S24293
chorion class CA protein mC11 precursor - silkworm
C:Species: Bombyx mori (silkworm)
C>Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C:Accession: S24293; A23548
R:Hitner, B.L.; Burke, W.D.; Eickbush, T.H.
Genetics 128, 595-606, 1991
A>Title: Sequence identity in an early chorion multigene family is the result of localiz
A:Reference number: S24291; MUID:91340109; PMID:1874417
A:Accession: S24293
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-121 <HIB>
A:Cross-references: UNIPROT:P08829; UNIPARC:UPI0000127788; EMBL:X58447; NID:G5854; PIDN:
R:lecanidou, R.; Rodakis, G.C.; Eickbush, T.H.; Kafatos, F.C.
Proc. Natl. Acad. Sci. U.S.A. 83, 6514-6518, 1986
A>Title: Evolution of the silk moth chorion gene superfamily: gene families CA and CB.
A:Reference number: A94122; MUID:86313609; PMID:3462711
A:Accession: A23548
A:Molecule type: mRNA
A:Residues: 16-121 <LEC>
A:Cross-references: UNIPARC:UPI00016888A; GB:M13834; NID:G155969; PIDN:AAA27828.1; PID:
C:Genetics:
A:introns: 19/3
C:Superfamily: chorion class A protein pc292
C:Keywords: egg

Query Match 65.9%; Score 29; DB 2; Length 121;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 109 CGPTCGC 115

RESULT 32
E82523
hypothetical protein XP2732 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: E82523
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: E82523
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-147 <SIM>
A:Cross-references: UNIPROT:Q9P978; UNIPARC:UPI00000C2B2A; GB:AE004079; GB:AE003849; NID:
R:experimental source: strain 9a5c
A:Simpon, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
as-Neto, E.; Docena, C.; El-Dorty, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kempfer, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.B.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshahko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XP2732

Query Match 65.9%; Score 29; DB 2; Length 147;

Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 138 CAPKRC 144

RESULT 33

T33130

Hypothetical protein C23H5.9 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C:Accession: T33130

R:Amam, E.; Kramer, J.

Submitted to the EMBL Data Library, May 1998

A:Description: The sequence of C. elegans cosmid C23H5.

A:Reference number: Z21286

A:Accession: T33130

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-163 <LAM>

A:Cross-references: UNIPROT:O61832; UNIPARC:UPI0000080484; EMBL:AF067609; PIDN:AAC17537.

A:Experimental source: strain Bristol N2; clone C23H5

C:Genetics:

A:Gene: CESP:C23H5.9

A:Map position: 4

A:Introns: 1/3; 101/3; 126/2

Query Match 65.9%; Score 29; DB 2; Length 163;
Best Local Similarity 57.1%; Pred. No. 1.6e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 44 CDPFGC 50

RESULT 34

P84680

Hypothetical protein At2g28090 [imported] - *Arabidopsis thaliana*C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C:Accession: P84680

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bent, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Ko, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Yencken, S.E.; Umayam, L.; Tallon, L.

Eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, D.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: P84680

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-245 <STO>

A:Cross-references: UNIPROT:Q9ZUV1; UNIPARC:UPI000009EFC4; GB:AE002093; NID:G4063749; PI

C:Genetics:

A:Gene: At2g28090

A:Map position: 2

Query Match 65.9%; Score 29; DB 2; Length 245;
Best Local Similarity 57.1%; Pred. No. 2.2e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 210 CQPGNC 216

RESULT 35

P90172

Hypothetical protein thif [imported] - *Sulfolobus solfataricus*C:Species: *Sulfolobus solfataricus*

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C:Accession: P90172

R:She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awey, M.J.; Chan-

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.

arrett, R.A.; Ragan, M.A.; Sersen, C.W.; Van der Oost, J.

Submitted to Genbank, April 2001

A:Description: *Sulfolobus solfataricus* complete genome.

A:Reference number: A99139

A:Accession: P90172

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-333 <KUR>

A:Cross-references: UNIPROT:Q980U4; UNIPARC:UPI0000064201; GB:AE006641; NID:G13813443; I

C:Genetics:

A:Gene: thif

Query Match 65.9%; Score 29; DB 2; Length 333;
Best Local Similarity 57.1%; Pred. No. 2.8e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 321 CYPSTGC 327

RESULT 36

T35248

probable oxidoreductase - *Streptomyces coelicolor*C:Species: *Streptomyces coelicolor*

C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C:Accession: T35248

R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.

Submitted to the EMBL Data Library, April 1999

A:Reference number: Z21573

A:Accession: T35248

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-348 <OLI>

A:Cross-references: UNIPROT:Q9X7M9; UNIPARC:UPI00000DAF9B; EMBL:AL049587; PIDN:CAB40673

A:Experimental source: strain A3 (2)

C:Genetics:

A:Gene: SCORDB:SCSF2A.06

Query Match 65.9%; Score 29; DB 2; Length 348;
Best Local Similarity 57.1%; Pred. No. 2.9e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 328 CYPDLC 334

RESULT 37

A83636

Hypothetical protein PA0063 [imported] - *Pseudomonas aeruginosa* (strain PA01)C:Species: *Pseudomonas aeruginosa*

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C:Accession: A83636

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mitsuuchi, S.D.; Warren, P.; Hickey, M.J.; B.

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lardig, K.; Lim,

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathc

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: A83636

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-375 <STO>

A:Cross-references: UNIPROT:Q91769; UNIPARC:UPI00000C4F10; GB:AE004446; GB:AE004091; NII

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA0063

Query Match 65.9%; Score 29; DB 2; Length 375;

Best Local Similarity 57.1%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 136 CFPDAGC 142

RESULT 38

S14885

hypothetical protein 8 - yeast (*Pichia angusta*)C/Species: *Pichia angusta*

C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #ext_change 09-Jul-2004

C/Accession: S14885

R/Krutilina, A.; Sergina, S.A.; Tikhomirova, L.P.; Kryukov, V.M.

submitted to the EMBL Data Library, April 1991

A/Description: Nucleotide sequence of *Hansenula polymorpha* DNA region complementing DAK.

A/Reference number: S14854

A/Accession: S14885

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-379 <NRU>

A/Cross-references: UNIPROT:Q04326; UNIPARC:UPI00006AASF; EMBL:X58862; NID:92762; PIND:

Query Match 65.9%; Score 29; DB 2; Length 379;
Best Local Similarity 57.1%; Pred. No. 3.1e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 372 CRPTTAC 378

RESULT 39

E96676

hypothetical protein T23K8.15 [imported] - *Arabidopsis thaliana*C/Species: *Arabidopsis thaliana* (mouse-ear cross)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #ext_change 09-Jul-2004

C/Accession: E96676

R/Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.;

ansen, N.F.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luroe, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: E96676

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-449 <STO>

A/Cross-references: UNIPROT:Q9S9K4; UNIPARC:UPI0000AA2D4; GB:AE005173; NID:94646203; PI

C/Genetics:

A/Map position: 1

Query Match 65.9%; Score 29; DB 2; Length 449;
Best Local Similarity 57.1%; Pred. No. 3.5e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 148 CQPALGC 154

RESULT 40

T23375

hypothetical protein K06G5.1 - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 09-Jul-2004

C/Accession: T23375

R/Mortimore, B.

submitted to the EMBL Data Library, November 1996

A/Reference number: Z19733

A/Accession: T23375

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-507 <ML>

A/Cross-references: UNIPROT:Q9XUJ9; UNIPARC:UPI000007B186; EMBL:Z81565; PIND:CA04581.1;

A/Experimental source: clone K06G5

C/Genetics:

A/Map position: X

A/Introns: 14/1; 71/3; 103/2; 135/1; 199/3; 242/3; 268/2; 306/2; 377/1; 430/3; 461/2

Query Match 65.9%; Score 29; DB 2; Length 507;
Best Local Similarity 57.1%; Pred. No. 3.8e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 52 CYPDAGC 58

RESULT 41

D88991

protein apx-1 [imported] - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #ext_change 10-May-2001

C/Accession: D88991

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A/Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology

A/Reference number: A75000; MUID:99069613; PMID:9851916

A/Note: see website genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele

A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A/Accession: D88991

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-513 <STO>

A/Cross-references: UNIPARC:UPI00017A664; GB:chr_V; PIND:AAC6353.1; PID:93808343; GSPD

C/Genetics:

A/Map position: 5

Query Match 65.9%; Score 29; DB 2; Length 513;
Best Local Similarity 57.1%; Pred. No. 3.9e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 207 CUPRAGC 213

RESULT 42

JC7570

Delta-4 protein - human

C/Species: *Homo sapiens* (man)

C/Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #ext_change 09-Jul-2004

C/Accession: JC7570

R:Toneya, T.; Tahara, T.; Nagao, K.; Yamada, Y.; Yamamoto, T.; Osawa, M.; Miyatani, S.;

J. Biochem. 129, 27-34, 2001

A/Title: Molecular cloning of Delta-4, a new mouse and human notch ligand.

A/Reference number: JC7569; MUID: 21064937; PMID:11134954

A/Accession: JC7570

A/Molecule type: mRNA

A/Residues: 1-685 <YON>

A/Cross-references: UNIPROT:Q9NR61; UNIPARC:UPI000047825; DDBJ:AB043894

C/Comment: This protein, a member of the Notch family of proteins, is a transmembrane re

ates the Notch signaling, the growth or differentiation of vascular endothelial cells.

C/Genetics:

A/Map position: delta-4

C/Superfamily: delta-4 protein; EGF homology

C:Keywords: transmembrane protein

Query Match 65.9%; Score 29; DB 2; Length 685;
Best Local Similarity 57.1%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 253 CIPHNCG 259

RESULT 43

JC7569

Delta-4 protein - mouse

C:Species: Mus musculus (house mouse)

C>Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004

C:Accession: J07569

R:Yoneya, T.; Tahara, T.; Nagao, K.; Yamada, Y.; Yamamoto, T.; Osawa, M.; Miyatani, S.;

J. Biochem. 129, 27-34, 2001

A:Title: Molecular cloning of Delta-4, a new mouse and human notch ligand.

A:Reference number: JC7569; MUID: 21064937; PMID:11134954

A:Accession: JC7569

A:Molecule type: mRNA

A:Residues: 1-686 <YON>

A:Cross-references: UNIPROT:Q9DBU9; UNIPARC:UPI000000B286; DDBJ:AB043893

C:Comment: This protein, a member of the Notch family of proteins, is a transmembrane re

ates the Notch signaling, the growth or differentiation of vascular endothelial cells.

C:Genetics:

A:Gene: delta-4

C:Superfamily: delta-4 protein; EGF homology

Query Match 65.9%; Score 29; DB 2; Length 686;
Best Local Similarity 57.1%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 254 CIPHNCG 260

RESULT 44

T28787

hypothetical protein C41D11.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T28787

R:Gatung, S.; Magill, L.

submitted to the EMBL Data Library, May 1997

A:Description: The sequence of C. elegans cosmid C41D11.

A:Reference number: Z20522

A:Accession: T28787

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-753 <GAT>

A:Cross-references: UNIPARC:UPI000017B826; EMBL:AF003740; PIDD:AA048141.1; GSPDB:GN00019

A:Experimental source: strain Bristol NZ; clone C41D11

C:Genetics:

A:Gene: CESP:C41D11.5

A:Map position: 1

A:Introns: 53/2; 81/3; 117/1; 250/3; 274/2; 357/3; 443/2; 485/3; 544/3; 585/3; 637/2

Query Match 65.9%; Score 29; DB 2; Length 753;
Best Local Similarity 57.1%; Pred. No. 5.2e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 338 CEPATNC 344

RESULT 45

A27340

complement C7 precursor [validated] - human

C:Species: Homo sapiens (man)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: A27340; B53072; C53072

R:Disclipio, R.G.; Chakravarti, D.N.; Muller-Berhard, H.J.; Fey, G.H.

J. Biol. Chem. 263, 549-560, 1988

A:Title: The structure of human complement component C7 and the C5b-7 complex.

A:Reference number: A27340; MUID:88087145; PMID:3335508

A:Accession: A27340

A:Molecule type: mRNA

A:Residues: 1-843 <DIS>

A:Cross-references: UNIPROT:P10643; UNIPARC:UPI0000127C58; GB:J03507; NID:G179715; PIDD:

R:Hobart, M.J.; Fernie, B.; Disclipio, R.G.;

Biochemistry 32, 6198-6205, 1993

A:Title: Structure of the human C6 gene.

A:Reference number: A53072; MUID:93291175; PMID:8512929

A:Accession: B53072

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-15, 'P', 17-21, 'R', 22-143 <HOB1>

A:Cross-references: UNIPARC:UPI00001742CF; GB:X72190; NID:G312799; GB:X72191; NID:G31280

A:Note: sequence extracted from NCBI backbone (NCBIP:134085)

A:Note: this translation is not annotated in Genbank entries HSC7X2, HSC7X3, and HSC7X4,

A:Accession: C53072

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 584-596, 'TY', 599-624, 'N', 626-627, 'L', 629-690, 'L' <HOB2>

A:Cross-references: UNIPARC:UPI00001742D0; GB:X72193; NID:G312797; GB:X72194; NID:G31279

A:Note: sequence extracted from NCBI backbone (NCBIP:134087)

A:Note: this translation is not annotated in Genbank entries HSC7X13, and HSC7X14, relea

R:Hoisteenge, U.; Blommestein, M.; Hesse, D.; Furanek, A.; Mitroshchenko, O.

J. Biol. Chem. 274, 32786-32794, 1999

A:Title: The four terminal components of the complement system are C-mannosylated on mu

A:Reference number: A59363; MUID:20020247; PMID:10551839

A:Contents: annotation

A:Note: identification and location of C-mannosylation sites by mass-spectroscopy

C:Genetics:

A:Gene: GDB:C7

A:Cross-references: GDB:119046; OMIM:217070

A:Map position: 5p13-5p13

C:Complex: combines with complement C5b-6 complex to form the complement C5b-7 complex

C:Function:

A:Description: as a component of the complement C5b-8 complex, assists in the polymeriza

A:Pathway: complement pathway

C:Superfamily: complement c6; complement factor H repeat homology; EGF homology; LDL rec

C:Keywords: complement alternate pathway; complement pathway; cytolysis; glycoprotein; m

F:1-22/Domain: signal sequence status predicted <SIG>

F:23-843/Product: complement component C7 #status predicted <MAT>

F:26-80/Domain: thrombospondin type 1 repeat homology <THR1>

F:85-119/Domain: LDL receptor ligand-binding repeat homology <LDL>

F:455-486/Domain: EGF homology <EGF>

F:499-546/Domain: thrombospondin type 1 repeat homology <THR2>

F:571-626/Domain: complement factor H repeat homology <FH01>

F:631-688/Domain: complement factor H repeat homology <FH02>

F:636/Modified site: 2'-mannosyl-tryptophan (Trp) #status experimental

F:702-754/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:503-506/Modified site: 2'-mannosyl-tryptophan (Trp) (partial) #status experimental

Query Match 65.9%; Score 29; DB 1; Length 843;
Best Local Similarity 57.1%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 73 CEPXTGC 79

RESULT 46

T16840

hypothetical protein T10B10.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C:Accession: T16840

R:Geisel, C.

submitted to the EMBL Data Library, October 1995
A:Description: The sequence of C. elegans cosmid T10E10.
A:Reference number: Z18588
A:Accession: T16840
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1101 <GE1>
A:Cross-references: UNIPROT:Q22378; UNIPARC:UPI000017BB8F; EMBL:U39644; NID:G1049339; PI
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:T10E10.4
A:Introns: 93/2; 152/2; 191/3; 209/2; 283/3; 303/1; 399/3; 421/1; 440/1; 465/1; 547/3; 7

Query Match 65.9%; Score 29; DB 2; Length 1101;
Best Local Similarity 57.1%; Pred. No. 7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTCG 10
DB 443 CTPGLGC 449

RESULT 47
T02766
adhesin Wt-1 - Afellowyces dermatitidis
C:Species: Afellowyces dermatitidis
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T02766
R:Hogan, L.H.; Josvai, S.; Klein, B.S.
J. Biol. Chem. 270, 30725-30732, 1995
A:Title: Genomic cloning, characterization, and functional analysis of the major surface
A:Reference number: Z14730; PMID:96107236; PMID:8530512
A:Accession: T02766
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1146 <HOG>
A:Cross-references: UNIPROT:O00048; UNIPARC:UPI000006C854; EMBL:U37772; NID:G1022895; PI
A:Experimental source: ATCC 26199

Query Match 65.9%; Score 29; DB 2; Length 1146;
Best Local Similarity 57.1%; Pred. No. 7.2e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTCG 10
DB 1067 CPTTSC 1073

RESULT 48
T26859
hypothetical protein Y43F8B.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T26859
R:Ainscough, R.
submitted to the EMBL Data Library, October 1998
A:Reference number: Z20278
A:Accession: T26859
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1743 <WIL>
A:Cross-references: UNIPROT:Q9XW5; UNIPARC:UPI0000078FD6; EMBL:AL032623; PIDN:CAA21511.
A:Experimental source: clone Y43F8B
C:Genetics:
A:Gene: CESP:Y43F8B.3
A:Introns: 65/1; 92/2; 128/1; 229/1; 367/1; 422/1; 486/2; 523/1; 571/1; 628/1; 857/2; 94

Query Match 65.9%; Score 29; DB 2; Length 1743;
Best Local Similarity 57.1%; Pred. No. 9.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTCG 10

DB 438 CAPGCGC 444

RESULT 49
T10053
laminin alpha 5 chain - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T10053
R:Miner, J.H.; Lewis, R.M.; Sanes, J.R.
submitted to the EMBL Data Library, November 1997
A:Reference number: Z16923
A:Accession: T10053
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3635 <MIN>
A:Cross-references: UNIPROT:O61001; UNIPARC:UPI000004C5E8; EMBL:U37501; NID:G2599231; PI
F:1942-1970/Domain: EGF homology <EGF>

Query Match 65.9%; Score 29; DB 2; Length 3635;
Best Local Similarity 57.1%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTCG 10
DB 1441 CHPLVGC 1447

RESULT 50
S25773
testis-specific protein Met84Db - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C:Accession: S25773; B56565
R:Kuhn, R.; Kuhn, C.; Boersch, D.; Glaetzer, K.H.; Schaefer, U.; Schaefer, M.
Mech. Dev. 35, 143-151, 1991
A:Title: A cluster of four genes selectively expressed in the male germ line of Drosophi
A:Reference number: A56565; PMID:92102953; PMID:1684716
A:Accession: S25773
A:Molecule type: DNA
A:Residues: 1-74 <KUH>
A:Cross-references: UNIPROT:Q01643; UNIPARC:UPI000012EAB6; EMBL:X67703; NID:G11072; PIDN
A:Note: sequence extracted from NCBI backbone (NCBIN:74217, NCBI:P.74220)
C:Genetics:
A:Gene: Met84Db
A:Cross-references: FlyBase:FBgn0004173
A:Map position: 3
C:Superfamily: fruit fly testis-specific protein
C:Keywords: spermatogenesis; tandem repeat

Query Match 63.6%; Score 28; DB 2; Length 74;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTCG 10
DB 68 CRPYCGC 74

Search completed: January 4, 2006, 16:09:43
Job time : 39.4522 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 ; Search time 88.9652 Seconds
(without alignments)
64.204 Million cell updates/sec

Title: US-09-932-322-1
Perfect score: 44
Sequence: 1 XXXCXPXTCXXX 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21.*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	81.8	192	8	ADs23148 Bacterial
2	35	79.5	79	4	AAU49413 Propionib
3	35	79.5	79	6	ABM45932 Propionib
4	35	79.5	180	7	ABO74428 Pseudomon
5	35	79.5	371	8	ADT04024 Human pro
6	34	77.3	7	3	AAy62764 PB-cadher
7	34	77.3	7	3	AAy61489 Cadherin-
8	34	77.3	7	3	AAy62007 Cadherin-
9	34	77.3	7	3	AAy62224 Cadherin-
10	34	77.3	7	5	ABU00550 B lymphoc
11	34	77.3	7	5	ABG33862 B lymphoc
12	34	77.3	13	5	ABJ00545 B lymphoc
13	34	77.3	13	5	ABG33861 B lymphoc
14	34	77.3	53	4	AAU44932 Propionib
15	34	77.3	53	6	ABM41451 Propionib
16	34	77.3	99	8	ADx87568 Plant full
17	34	77.3	113	6	ADA54961 Human pro
18	34	77.3	138	3	AAAB42027 Human ORF
19	34	77.3	250	7	ABO75329 Pseudomon
20	34	77.3	312	5	ABB93198 Hericida
21	34	77.3	532	8	ADN17632 Bacterial
22	34	77.3	540	7	ADT05702 Bacterial
23	34	77.3	3226	4	ABG28408 Novel hum
24	33	75.0	158	4	AAU40006 Propionib

25	33	75.0	158	6	ABM36525
26	33	75.0	239	8	ADT57626
27	32	72.7	751	3	AAg43626
28	32	72.7	798	3	AAg43625
29	32	72.7	835	7	AAg43624
30	31	70.5	90	7	AAy63264
31	31	70.5	114	5	ABO81251
32	31	70.5	148	7	ABP06368
33	31	70.5	183	7	ADJ71040
34	31	70.5	193	4	ADP58925
35	31	70.5	193	4	AAO00015
36	31	70.5	193	7	ADP60298
37	31	70.5	238	6	ABP92529
38	31	70.5	239	6	ABP99348
39	31	70.5	239	6	ABP67713
40	31	70.5	285	8	ADx95230
41	31	70.5	310	7	ABO74638
42	31	70.5	354	8	ADU17203
43	31	70.5	444	7	ABO70819
44	31	70.5	482	5	ABB93340
45	31	70.5	580	2	AAW32363
46	31	70.5	580	2	AAW32431
47	31	70.5	580	2	AAW64303
48	31	70.5	580	2	AAW81666
49	31	70.5	580	2	AAW38968
50	31	70.5	580	2	AAW39105
51	31	70.5	602	2	AAW04996
52	31	70.5	602	8	ADU17226
53	31	70.5	602	2	AAW04992
54	31	70.5	804	7	ABO75305
55	31	70.5	19938	6	ABB98398
56	30	68.2	7	3	AAy64260
57	30	68.2	13	9	ADV77018
58	30	68.2	55	9	ADV77016
59	30	68.2	63	5	ABP33508
60	30	68.2	76	4	ABH16814
61	30	68.2	87	4	AAW78641
62	30	68.2	94	4	AAU22004
63	30	68.2	94	7	ADB45972
64	30	68.2	94	8	ADJ07390
65	30	68.2	96	4	ABG04785
66	30	68.2	109	7	ABO74041
67	30	68.2	121	7	ABO76722
68	30	68.2	126	7	ABO72980
69	30	68.2	133	7	ADB99069
70	30	68.2	133	7	ADB82537
71	30	68.2	133	7	ABO75949
72	30	68.2	142	4	ADM19798
73	30	68.2	175	4	AAU72022
74	30	68.2	175	4	AAU72525
75	30	68.2	187	4	AAU61592
76	30	68.2	187	6	ABM58111
77	30	68.2	208	8	ADY13040
78	30	68.2	208	8	AAK51896
79	30	68.2	211	2	AAK51897
80	30	68.2	342	2	AAV42780
81	30	68.2	345	2	AAW13907
82	30	68.2	366	4	ABG15722
83	30	68.2	391	4	ABG20618
84	30	68.2	451	9	ABM94980
85	30	68.2	452	9	ADM26287
86	30	68.2	459	8	ADRO9703
87	30	68.2	469	8	ADM90874
88	30	68.2	494	8	ADM90873
89	30	68.2	547	4	ABG08052
90	30	68.2	575	8	ADx94751
91	30	68.2	613	9	ABM90694
92	30	68.2	638	9	ABG20620
93	30	68.2	672	4	ABG14072
94	30	68.2	685	4	ABG29714
95	30	68.2	698	7	ADM05627
96	30	68.2	718	6	ADA33353
97	30	68.2	734	8	ADO28637

ABM36525	Propionib
ADT57626	Plant pol
AAg43626	Arabidops
AAg43625	Arabidops
AAg43624	Arabidops
AAy63264	Protocad
ABO81251	Pseudomon
ABP06368	Human ORF
ADJ71040	Human hea
ADP58925	Human pol
AAO00015	Human pol
ADP60298	Human con
ABP92529	Orthosomy
ABP99348	Orthosomy
ABP67713	Streptomy
ADx95230	Plant full
ABO74638	Pseudomon
ADU17203	M. tuberc
ABO70819	Pseudomon
ABB93340	Herbicida
AAW32363	Mycobacte
AAW32431	Mycobacte
AAW64303	Mycobacte
AAW81666	M. tuberc
AAW38968	M. tuberc
AAW39105	M. tuberc
AAW04996	Mycobacte
ADU17226	M. tuberc
AAW04992	Mycobacte
ABO75305	Pseudomon
ABB98398	Streptomy
AAy64260	Cadherin-
ADV77018	Human TAA
ADV77016	Human TAA
ABP33508	Human ORF
ABH16814	Human ner
AAW78641	Human pro
AAU22004	Human car
ADB45972	Human car
ADJ07390	Human car
ABG04785	Novel hum
ABO74041	Pseudomon
ABO76722	Pseudomon
ABO72980	Pseudomon
ADB99069	LRPS cons
ADB82537	LRPS pept
ABO75949	Pseudomon
ADM19798	Protein e
AAU72022	Human olf
AAU72525	Human OR-
AAU61592	Propionib
ABM58111	Propionib
ADY13040	Plant full
AAK51896	Chitin bi
AAK51897	Chitin bi
AAV42780	Rat neuro
AAW13907	Serpin pr
ABG15722	Novel hum
ABG20618	Novel hum
ABM94980	M. xanthu
ADM26287	Hypether
ADT09703	Human pro
ADM90874	Human pha
ADM90873	Human pha
ADx94751	Novel hum
ABG08052	Novel hum
ADx94751	Plant full
ABM90694	M. xanthu
ABG20620	Novel hum
ABG14072	Novel hum
ABG29714	Novel hum
ADM05627	Human pro
ADA33353	Acinetoba
ADO28637	Human ADA

98	30	68.2	764	5	ABR05596	AbD05596	Human tes	171	29	65.9	258	4	ABG27045	Abg27045	Novel hum
99	30	68.2	767	4	ABG11534	AbG11534	Novel hum	172	29	65.9	263	7	ABO68160	ABO68160	Pseudomon
100	30	68.2	767	7	ADP60558	AdP60558	Human con	173	29	65.9	293	2	AAR28148	AAR28148	Sugar bee
101	30	68.2	789	7	ABR55542	ABr55542	Human pro	174	29	65.9	336	8	ADO6304	ADO6304	Novel hum
102	30	68.2	834	8	ADN47192	AdN47192	Thermococ	175	29	65.9	342	7	ABO63151	ABO63151	Klebsiell
103	30	68.2	995	7	ABM89334	ABm89334	Rice abio	176	29	65.9	359	8	ADY22332	ADY22332	Plant ful
104	30	68.2	2161	2	AAW87724	AAw87724	Rupesctris	177	29	65.9	361	4	AAU30935	AAU30935	Novel hum
105	29	65.9	7	9	ABR63928	ABr63928	Lysozyme	178	29	65.9	394	5	ABP69774	ABP69774	Human pol
106	29	65.9	10	4	AA96135	AA96135	Human com	179	29	65.9	429	4	ABR65487	ABR65487	Drosophil
107	29	65.9	10	4	AA966045	AA966045	Human com	180	29	65.9	436	7	ABO71510	ABO71510	Pseudomon
108	29	65.9	10	4	AA966075	AA966075	Human com	181	29	65.9	436	9	ABR15465	ABR15465	Human pol
109	29	65.9	10	4	AA966083	AA966083	Human com	182	29	65.9	439	7	ABO73150	ABO73150	Pseudomon
110	29	65.9	10	4	AA966073	AA966073	Human com	183	29	65.9	455	5	ABR91454	ABR91454	Herpidea
111	29	65.9	10	4	AA966081	AA966081	Human com	184	29	65.9	485	6	ABO01387	ABO01387	Human pro
112	29	65.9	12	3	AA826019	AA826019	Human IGE	185	29	65.9	485	8	ADN96146	ADN96146	Human NOV
113	29	65.9	12	3	AAU16744	AAu16744	Peptide E	186	29	65.9	489	6	ABO01396	ABO01396	Human pro
114	29	65.9	12	5	ABU00367	ABj00367	Human IGE	187	29	65.9	489	8	ADN96164	ADN96164	Human NOV
115	29	65.9	18	4	AAW65407	AAw65407	Human bra	188	29	65.9	500	2	AAW94496	AAW94496	Human del
116	29	65.9	19	4	AA673162	AA673162	Protease	189	29	65.9	503	3	AAV68936	AAV68936	Amino aci
117	29	65.9	19	4	AA673155	AA673155	Protease	190	29	65.9	504	3	AAV68935	AAV68935	Amino aci
118	29	65.9	50	4	AAU62562	AAu62562	Protoniob	191	29	65.9	505	3	AAV68934	AAV68934	Amino aci
119	29	65.9	50	6	ABM59081	ABm59081	Protoniob	192	29	65.9	505	3	AAV68933	AAV68933	Amino aci
120	29	65.9	50	6	ABO59332	ABo59332	Human gen	193	29	65.9	506	3	AAV68933	AAV68933	Amino aci
121	29	65.9	53	4	AAU49306	AAu49306	Protoniob	194	29	65.9	506	3	AAV68917	AAV68917	Amino aci
122	29	65.9	53	4	ABM45825	ABm45825	Protoniob	195	29	65.9	507	3	AAV68932	AAV68932	Amino aci
123	29	65.9	55	3	AA842800	AA842800	Human ORF	196	29	65.9	507	3	AAV68916	AAV68916	Amino aci
124	29	65.9	64	4	ABG16398	ABg16398	Novel hum	197	29	65.9	508	3	AAV68915	AAV68915	Amino aci
125	29	65.9	71	5	ABP00915	ABp00915	Human ORF	198	29	65.9	508	3	AAV68931	AAV68931	Amino aci
126	29	65.9	71	7	ABO82044	ABo82044	Pseudomon	199	29	65.9	509	3	AAV68914	AAV68914	Amino aci
127	29	65.9	85	4	AAW65401	AAw65401	Human bra	200	29	65.9	510	3	AAV68913	AAV68913	Amino aci
128	29	65.9	85	6	ADA56925	Ada56925	Human sec	201	29	65.9	524	3	AAV68942	AAV68942	Amino aci
129	29	65.9	85	6	ADA40775	Ada40775	Human sec	202	29	65.9	525	3	AAV68941	AAV68941	Amino aci
130	29	65.9	85	7	ADC74135	AdC74135	Human sec	203	29	65.9	526	3	AAV68924	AAV68924	Amino aci
131	29	65.9	86	3	AAV91452	AAv91452	Human sec	204	29	65.9	526	3	AAV68924	AAV68924	Amino aci
132	29	65.9	86	8	ADL71524	ADl71524	Novel hum	205	29	65.9	527	3	AAV68923	AAV68923	Amino aci
133	29	65.9	92	4	ABR03747	ABr03747	Human mus	206	29	65.9	527	3	AAV68937	AAV68937	Amino aci
134	29	65.9	92	6	ABU13041	ABu13041	Novel hum	207	29	65.9	528	3	AAV68922	AAV68922	Amino aci
135	29	65.9	104	3	AA600170	AA600170	Human sec	208	29	65.9	528	6	ABR39450	ABR39450	Human bol
136	29	65.9	104	3	AAW48349	AAw48349	Human bre	209	29	65.9	528	3	AAV68921	AAV68921	Amino aci
137	29	65.9	106	2	AAW48349	AAw48349	Novel hum	210	29	65.9	529	3	AAV68937	AAV68937	Amino aci
138	29	65.9	109	4	ABG26326	ABg26326	Novel hum	211	29	65.9	529	3	AAV68937	AAV68937	Amino aci
139	29	65.9	112	4	AAU40576	AAu40576	Protoniob	212	29	65.9	530	3	AAV68920	AAV68920	Amino aci
140	29	65.9	112	6	ABM73095	ABm73095	Protoniob	213	29	65.9	531	4	AAV68919	AAV68919	Amino aci
141	29	65.9	116	4	AAW73965	AAw73965	Human bon	214	29	65.9	534	5	ABG11485	ABG11485	D. melano
142	29	65.9	116	5	ABG43858	ABg43858	Human pep	215	29	65.9	605	4	ABG14995	ABG14995	Novel hum
143	29	65.9	116	7	ABO80121	ABo80121	Pseudomon	216	29	65.9	621	6	AAE34037	AAE34037	Murine no
144	29	65.9	122	4	AAU51053	AAu51053	Protoniob	217	29	65.9	653	8	ADP88277	ADP88277	Human del
145	29	65.9	122	6	ABM47572	ABm47572	Protoniob	218	29	65.9	659	2	AAW94497	AAW94497	Human del
146	29	65.9	126	2	AAV59997	AAv59997	Human end	219	29	65.9	659	3	AAV68930	AAV68930	Amino aci
147	29	65.9	129	8	ADP09225	ADp09225	Human pro	220	29	65.9	660	3	AAV68929	AAV68929	Amino aci
148	29	65.9	134	6	ADP06446	ADp06446	Alloiococ	221	29	65.9	661	3	AAV68928	AAV68928	Amino aci
149	29	65.9	134	7	ADP58881	ADp58881	Human pol	222	29	65.9	662	3	AAV68927	AAV68927	Amino aci
150	29	65.9	137	6	ABO01386	ABo01386	Human pro	223	29	65.9	663	3	AAV68926	AAV68926	Amino aci
151	29	65.9	137	8	ADN96144	ADN96144	Human NOV	224	29	65.9	664	3	AAV68925	AAV68925	Amino aci
152	29	65.9	142	4	AAU41380	AAu41380	Protoniob	225	29	65.9	677	3	AAV68948	AAV68948	Amino aci
153	29	65.9	142	6	ABM73899	ABm73899	Protoniob	226	29	65.9	678	3	AAV68947	AAV68947	Amino aci
154	29	65.9	146	6	ADP06448	ADp06448	Alloiococ	227	29	65.9	679	3	AAV68946	AAV68946	Amino aci
155	29	65.9	151	8	ADY06118	ADy06118	Plant ful	228	29	65.9	680	3	AAV68945	AAV68945	Amino aci
156	29	65.9	155	4	ABG26325	ABg26325	Novel hum	229	29	65.9	681	3	AAV68944	AAV68944	Amino aci
157	29	65.9	159	7	ABO78806	ABo78806	Pseudomon	230	29	65.9	682	3	AAV68943	AAV68943	Amino aci
158	29	65.9	159	7	ABO78806	ABo78806	Pseudomon	231	29	65.9	685	2	AAW80813	AAW80813	Nucleotid
159	29	65.9	161	4	ABG04542	ABg04542	Novel hum	232	29	65.9	685	2	AAW94507	AAW94507	Human del
160	29	65.9	167	5	ADP41479	ADp41479	Human CD-	233	29	65.9	685	3	AAV76912	AAV76912	Human not
161	29	65.9	174	4	AAU16948	AAu16948	Human nov	234	29	65.9	685	4	AAU12215	AAU12215	Amino aci
162	29	65.9	188	4	AAU18112	AAu18112	Novel hum	235	29	65.9	685	4	AAU12215	AAU12215	Amino aci
163	29	65.9	188	4	AAU17021	AAu17021	Human nov	236	29	65.9	685	5	AAU83698	AAU83698	Human PRO
164	29	65.9	188	4	ABR10465	ABr10465	Human cDN	237	29	65.9	685	5	ABR84954	ABR84954	Human PRO
165	29	65.9	188	4	AAU19912	AAu19912	Novel hum	238	29	65.9	685	5	ABR95560	ABR95560	Human ang
166	29	65.9	188	5	ABR05739	ABr05739	Novel hum	239	29	65.9	685	5	ADY31954	ADY31954	Novel hum
167	29	65.9	188	5	ABP67052	ABp67052	Human pol	240	29	65.9	685	6	ABO17659	ABO17659	Novel hum
168	29	65.9	197	7	ABM89228	ABm89228	Rice abio	241	29	65.9	685	6	ABU80845	ABU80845	Human PRO
169	29	65.9	220	8	ABO59256	ABo59256	Human gen	242	29	65.9	685	6	ABO33811	ABO33811	Novel hum
170	29	65.9	235	9	ABM95832	ABm95832	M. xanthu	243	29	65.9	685	6	ABU80913	ABU80913	Human PRO

244	29	65.9	685	6	ABP97826	Abp97826	Amino aci	317	29	65.9	685	7	ADB22570	Adb22570	Human	PRO
245	29	65.9	685	6	ABR39449	Abri39449	Human H4-	318	29	65.9	685	7	ADB23343	Adb23343	Human	PRO
246	29	65.9	685	6	ABU66613	Abu66613	Human PRO	319	29	65.9	685	7	ADB29205	Adb29205	Human	PRO
247	29	65.9	685	6	ABU55880	Abu55880	Human not	320	29	65.9	685	7	ADB15128	Adb15128	Human	PRO
248	29	65.9	685	6	AA534036	Aa534036	Human not	321	29	65.9	685	7	ADB83704	Adb83704	Human	PRO
249	29	65.9	685	6	ABU59694	Abu59694	Novel sec	322	29	65.9	685	7	ADB80810	Adb80810	Novel	hum
250	29	65.9	685	6	ABO24884	AbO24884	Human sec	323	29	65.9	685	7	ADB73351	Adb73351	Novel	hum
251	29	65.9	685	6	ABP72568	Abp72568	Human Not	324	29	65.9	685	7	ADB83830	Adb83830	Novel	hum
252	29	65.9	685	6	ABU82154	Abu82154	Novel hum	325	29	65.9	685	7	ADB78433	Adb78433	Novel	hum
253	29	65.9	685	6	ABU66889	Abu66889	Human sec	326	29	65.9	685	7	ADB37828	Adb37828	Novel	hum
254	29	65.9	685	6	ADA55607	Ada55607	Novel hum	327	29	65.9	685	7	ADB66300	Adb66300	Novel	hum
255	29	65.9	685	6	ADA76038	Ada76038	Human PRO	328	29	65.9	685	7	ADB85081	Adb85081	Human	PRO
256	29	65.9	685	6	ABJ72334	Abj72334	Human PRO	329	29	65.9	685	7	ADB89380	Adb89380	Human	PRO
257	29	65.9	685	6	ADA18668	Ada18668	Human PRO	330	29	65.9	685	7	ADB90112	Adb90112	Human	PRO
258	29	65.9	685	6	ADA61311	Ada61311	Homo sapi	331	29	65.9	685	7	ADB39213	Adb39213	Novel	hum
259	29	65.9	685	6	ADB19096	Adb19096	Novel hum	332	29	65.9	685	7	ADB78187	Adb78187	Novel	hum
260	29	65.9	685	6	ADB27637	Adb27637	Human PRO	333	29	65.9	685	7	ADB87253	Adb87253	Human	PRO
261	29	65.9	685	6	ADB86116	Adb86116	Novel hum	334	29	65.9	685	7	ADB84835	Adb84835	Human	PRO
262	29	65.9	685	6	ADB15680	Adb15680	Human PRO	335	29	65.9	685	7	ADB46836	Adb46836	Novel	hum
263	29	65.9	685	6	ADA47466	Ada47466	Human PRO	336	29	65.9	685	7	ADB83950	Adb83950	Novel	hum
264	29	65.9	685	6	ADA67261	Ada67261	Human PRO	337	29	65.9	685	7	ADB86443	Adb86443	Human	PRO
265	29	65.9	685	6	ADB30268	Adb30268	Human PRO	338	29	65.9	685	7	ADB73105	Adb73105	Novel	hum
266	29	65.9	685	6	ADA85564	Ada85564	Novel hum	339	29	65.9	685	7	ADB77048	Adb77048	Novel	hum
267	29	65.9	685	6	ADA96776	Ada96776	Human PRO	340	29	65.9	685	7	ADB34205	Adb34205	Human	PRO
268	29	65.9	685	6	ADA79080	Ada79080	Human PRO	341	29	65.9	685	7	ADB35309	Adb35309	Human	PRO
269	29	65.9	685	6	ADA87219	Ada87219	Novel hum	342	29	65.9	685	7	ADB33653	Adb33653	Human	PRO
270	29	65.9	685	6	ADB16421	Adb16421	Human PRO	343	29	65.9	685	7	ADB34757	Adb34757	Human	PRO
271	29	65.9	685	6	ADA91513	Ada91513	Novel hum	344	29	65.9	685	7	ADB35681	Adb35681	Human	PRO
272	29	65.9	685	6	ADB14576	Adb14576	Human PRO	345	29	65.9	685	7	ADB46256	Adb46256	Novel	hum
273	29	65.9	685	6	ADB18537	Adb18537	Novel hum	346	29	65.9	685	7	ADC36943	Adc36943	Human	PRO
274	29	65.9	685	6	ADA93752	Ada93752	Human PRO	347	29	65.9	685	7	ADC21933	Adc21933	Human	PRO
275	29	65.9	685	6	ADB19648	Adb19648	Novel hum	348	29	65.9	685	7	ADC50129	Adc50129	Novel	hum
276	29	65.9	685	6	ADB12960	Adb12960	Human PRO	349	29	65.9	685	7	ADC71676	Adc71676	Novel	hum
277	29	65.9	685	6	ABO43192	AbO43192	Novel hum	350	29	65.9	685	7	ADC59655	Adc59655	Novel	hum
278	29	65.9	685	6	ADA74214	Ada74214	Human PRO	351	29	65.9	685	7	ADC49964	Adc49964	Novel	hum
279	29	65.9	685	6	ADB24447	Adb24447	Human PRO	352	29	65.9	685	7	ADC49163	Adc49163	Novel	hum
280	29	65.9	685	6	ADA81971	Ada81971	Human PRO	353	29	65.9	685	7	ADC49680	Adc49680	Novel	hum
281	29	65.9	685	6	ADA74934	Ada74934	Human PRO	354	29	65.9	685	7	ADC47541	Adc47541	Novel	hum
282	29	65.9	685	6	ADA85012	Ada85012	Novel hum	355	29	65.9	685	7	ADC52662	Adc52662	Novel	hum
283	29	65.9	685	6	ADA84460	Ada84460	Novel hum	356	29	65.9	685	7	ADC57016	Adc57016	Novel	hum
284	29	65.9	685	6	ADB29716	Adb29716	Human PRO	357	29	65.9	685	7	ADC60207	Adc60207	Novel	hum
285	29	65.9	685	6	ADA80244	Ada80244	Human PRO	358	29	65.9	685	7	ADC50682	Adc50682	Novel	hum
286	29	65.9	685	6	ADA75486	Ada75486	Human PRO	359	29	65.9	685	7	ADC65209	Adc65209	Human	PRO
287	29	65.9	685	6	ADA67711	Ada67711	Human PRO	360	29	65.9	685	7	ADC54307	Adc54307	Novel	hum
288	29	65.9	685	6	ADB35007	Adb35007	Human PRO	361	29	65.9	685	7	ADC53268	Adc53268	Novel	hum
289	29	65.9	685	6	ADA93183	Ada93183	Human PRO	362	29	65.9	685	7	ADC58791	Adc58791	Novel	hum
290	29	65.9	685	6	ADB26533	Adb26533	Human PRO	363	29	65.9	685	7	ADC55669	Adc55669	Novel	hum
291	29	65.9	685	6	ADB30820	Adb30820	Human PRO	364	29	65.9	685	7	ADC58239	Adc58239	Novel	hum
292	29	65.9	685	6	ABR61827	Abri61827	Human del	365	29	65.9	685	7	ADC47286	Adc47286	Novel	hum
293	29	65.9	685	6	ABJ72462	Abj72462	Human PRO	366	29	65.9	685	7	ADD02913	AdD02913	Novel	hum
294	29	65.9	685	6	ADA60748	Ada60748	Homo Sapi	367	29	65.9	685	7	ADC89905	Adc89905	Novel	hum
295	29	65.9	685	6	ADB23895	Adb23895	Human PRO	368	29	65.9	685	7	ADC69324	Adc69324	Human	PRO
296	29	65.9	685	6	ADA96224	Ada96224	Human PRO	369	29	65.9	685	7	ADC48213	Adc48213	Human	PRO
297	29	65.9	685	6	ADA80796	Ada80796	Human PRO	370	29	65.9	685	7	ADD09742	AdD09742	Human	PRO
298	29	65.9	685	6	ADA95672	Ada95672	Human PRO	371	29	65.9	685	7	ADC78161	Adc78161	Novel	hum
299	29	65.9	685	6	ADB25981	Adb25981	Human PRO	372	29	65.9	685	7	ADD04317	AdD04317	Novel	hum
300	29	65.9	685	6	ADB21466	Adb21466	Novel hum	373	29	65.9	685	7	ADD06396	AdD06396	Novel	hum
301	29	65.9	685	6	ABO34357	AbO34357	Human sec	374	29	65.9	685	7	ADC80273	Adc80273	Novel	hum
302	29	65.9	685	7	ADA77245	Ada77245	Human PRO	375	29	65.9	685	7	ADD10780	AdD10780	Human	PRO
303	29	65.9	685	7	ADB17985	Adb17985	Human PRO	376	29	65.9	685	7	ADD10565	AdD10565	Human	PRO
304	29	65.9	685	7	ADA86668	Ada86668	Novel hum	377	29	65.9	685	7	ADC47661	Adc47661	Human	PRO
305	29	65.9	685	7	ADA87771	Ada87771	Novel hum	378	29	65.9	685	7	ADC77915	Adc77915	Novel	hum
306	29	65.9	685	7	ABR61756	Abri61756	Human del	379	29	65.9	685	7	ADC79721	Adc79721	Novel	hum
307	29	65.9	685	7	ADA46159	Ada46159	Novel hum	380	29	65.9	685	7	ADD11525	AdD11525	Human	sec
308	29	65.9	685	7	ADB28189	Adb28189	Human PRO	381	29	65.9	685	7	ADD09190	AdD09190	Human	PRO
309	29	65.9	685	7	ADB28741	Adb28741	Human PRO	382	29	65.9	685	7	ADD50878	AdD50878	Novel	hum
310	29	65.9	685	7	ADA76693	Ada76693	Human PRO	383	29	65.9	685	7	ADD40903	AdD40903	Novel	hum
311	29	65.9	685	7	ADA88323	Ada88323	Novel hum	384	29	65.9	685	7	ADD52042	AdD52042	Human	PRO
312	29	65.9	685	7	ADA97328	Ada97328	Human PRO	385	29	65.9	685	7	ADD51124	AdD51124	Novel	hum
313	29	65.9	685	7	ADB27085	Adb27085	Human PRO	386	29	65.9	685	7	ADD52782	AdD52782	Human	PRO
314	29	65.9	685	7	ADB22018	Adb22018	Novel hum	387	29	65.9	685	7	ADD53334	AdD53334	Novel	hum
315	29	65.9	685	7	ABJ72164	Abj72164	Human mem	388	29	65.9	685	7	ADD37318	AdD37318	Human	sec
	29	65.9	685	7	ADA66709	Ada66709	Human PRO	389	29	65.9	685	7	ADD51490	AdD51490	Human	PRO

390	29	65.9	685	7	ADD02289	Ad002289	Human	PRO	463	29	65.9	685	8	ADE41526	Ad041526	Human	sec
391	29	65.9	685	7	ADD50605	Ad050605	Human	PRO	464	29	65.9	685	8	ADE23065	Ad023065	Human	PRO
392	29	65.9	685	7	ADD01123	Ad001123	Human	PRO	465	29	65.9	685	8	ADE21334	Ad021334	Novel	hum
393	29	65.9	685	7	ADD53905	Ad053905	Novel	hum	466	29	65.9	685	8	ADD77449	Ad077449	Novel	hum
394	29	65.9	685	7	ADD50359	Ad050359	Human	PRO	467	29	65.9	685	8	ADE20596	Ad020596	Novel	hum
395	29	65.9	685	7	ADD50359	Ad050359	Novel	hum	468	29	65.9	685	8	ADD75661	Ad075661	Human	PRO
396	29	65.9	685	7	ADD92222	Ad092222	Human	PRO	469	29	65.9	685	8	ADD74423	Ad074423	Human	PRO
397	29	65.9	685	7	ADD92222	Ad092222	Human	PRO	470	29	65.9	685	8	ADD74423	Ad074423	Human	PRO
398	29	65.9	685	7	ADD91118	Ad091118	Human	PRO	471	29	65.9	685	8	ADD76153	Ad076153	Novel	hum
399	29	65.9	685	7	ADD32029	Ad032029	Novel	hum	472	29	65.9	685	8	ADD85645	Ad085645	Novel	hum
400	29	65.9	685	7	ADD21961	Ad021961	Human	PRO	473	29	65.9	685	8	AD023617	Ad023617	Human	PRO
401	29	65.9	685	7	ADD79185	Ad079185	Human	PRO	474	29	65.9	685	8	ADE24260	Ad024260	Human	PRO
402	29	65.9	685	7	AD041721	Ad041721	Human	PRO	475	29	65.9	685	8	ADD87085	Ad087085	Human	PRO
403	29	65.9	685	7	AD017538	Ad017538	Human	PRO	476	29	65.9	685	8	ADE05194	Ad0505194	Human	PRO
404	29	65.9	685	7	AD091670	Ad091670	Human	PRO	477	29	65.9	685	8	ADD75407	Ad075407	Human	PRO
405	29	65.9	685	7	AD033133	Ad033133	Novel	hum	478	29	65.9	685	8	ADD76951	Ad076951	Novel	hum
406	29	65.9	685	7	AD033685	Ad033685	Novel	hum	479	29	65.9	685	8	ADD86719	Ad086719	Novel	hum
407	29	65.9	685	7	AD079737	Ad079737	Human	PRO	480	29	65.9	685	8	ADE88851	Ad088851	Human	PRO
408	29	65.9	685	7	ADD92774	Ad092774	Human	PRO	481	29	65.9	685	8	ADD78187	Ad078187	Novel	hum
409	29	65.9	685	7	AD019194	Ad019194	Human	PRO	482	29	65.9	685	8	ADE18090	Ad018090	Human	PRO
410	29	65.9	685	7	AD018642	Ad018642	Human	PRO	483	29	65.9	685	8	ADE88399	Ad088399	Human	PRO
411	29	65.9	685	7	AD042838	Ad042838	Human	PRO	484	29	65.9	685	8	ADD77695	Ad077695	Novel	hum
412	29	65.9	685	7	ADD95627	Ad095627	Human	PRO	485	29	65.9	685	8	ADD77941	Ad077941	Novel	hum
413	29	65.9	685	7	AD022513	Ad022513	Human	PRO	486	29	65.9	685	8	ADD85399	Ad085399	Novel	hum
414	29	65.9	685	7	AD078631	Ad078631	Human	PRO	487	29	65.9	685	8	ADD73931	Ad073931	Human	PRO
415	29	65.9	685	7	AD032581	Ad032581	Novel	hum	488	29	65.9	685	8	ADD74669	Ad074669	Human	PRO
416	29	65.9	685	7	AD042273	Ad042273	Human	PRO	489	29	65.9	685	8	ADD77197	Ad077197	Novel	hum
417	29	65.9	685	7	ADD80289	Ad080289	Human	PRO	490	29	65.9	685	8	ADD85891	Ad085891	Novel	hum
418	29	65.9	685	7	ADD89317	Ad089317	Human	PRO	491	29	65.9	685	8	ADE05540	Ad0505540	Human	PRO
419	29	65.9	685	7	AD040601	Ad040601	Human	PRO	492	29	65.9	685	8	ADD74915	Ad074915	Human	PRO
420	29	65.9	685	7	AD040400	Ad040400	Human	PRO	493	29	65.9	685	8	ADE94419	Ad094419	Human	PRO
421	29	65.9	685	7	AD032529	Ad032529	Human	PRO	494	29	65.9	685	8	ADE90830	Ad090830	Human	PRO
422	29	65.9	685	7	AD063370	Ad063370	Human	Not	495	29	65.9	685	8	ADE94971	Ad094971	Human	PRO
423	29	65.9	685	7	AD021238	Ad021238	Novel	hum	496	29	65.9	685	8	ADE93081	Ad093081	Human	PRO
424	29	65.9	685	7	AD032879	Ad032879	Novel	hum	497	29	65.9	685	8	AD034662	Ad034662	Human	PRO
425	29	65.9	685	7	AD097214	Ad097214	Human	PRO	498	29	65.9	685	8	ADE91977	Ad091977	Novel	hum
426	29	65.9	685	7	AD027922	Ad027922	Human	del	499	29	65.9	685	8	ADE90278	Ad090278	Human	PRO
427	29	65.9	685	7	AD027866	Ad027866	Human	del	500	29	65.9	685	8	ADE91425	Ad091425	Human	PRO
428	29	65.9	685	7	AD027933	Ad027933	Human	del	501	29	65.9	685	8	AD0505727	Ad0505727	Novel	hum
429	29	65.9	685	7	AD027899	Ad027899	Human	del	502	29	65.9	685	8	AD027281	Ad027281	Human	PRO
430	29	65.9	685	7	AD027944	Ad027944	Human	del	503	29	65.9	685	8	AD020204	Ad020204	Human	PRO
431	29	65.9	685	7	AD027911	Ad027911	Human	del	504	29	65.9	685	8	AD021790	Ad021790	Novel	hum
432	29	65.9	685	7	AD080278	Ad080278	Human	PRO	505	29	65.9	685	8	AD019860	Ad019860	Human	PRO
433	29	65.9	685	7	AD079726	Ad079726	Human	PRO	506	29	65.9	685	8	AD097766	Ad097766	Human	PRO
434	29	65.9	685	7	AD055018	Ad055018	Novel	hum	507	29	65.9	685	8	AD023983	Ad023983	Novel	hum
435	29	65.9	685	7	AD055570	Ad055570	Novel	hum	508	29	65.9	685	8	AD098337	Ad098337	Human	PRO
436	29	65.9	685	7	AD0163789	Ad0163789	Novel	hum	509	29	65.9	685	8	AD031168	Ad031168	Human	PRO
437	29	65.9	685	7	AD016738	Ad016738	Novel	hum	510	29	65.9	685	8	AD098889	Ad098889	Human	PRO
438	29	65.9	685	7	AD0163237	Ad0163237	Novel	hum	511	29	65.9	685	8	AD016474	Ad016474	Human	PRO
439	29	65.9	685	7	AD081651	Ad081651	Novel	hum	512	29	65.9	685	8	AD004933	Ad004933	Human	PRO
440	29	65.9	685	7	AD081099	Ad081099	Novel	hum	513	29	65.9	685	8	AD019200	Ad019200	Human	PRO
441	29	65.9	685	7	AD082268	Ad082268	Novel	hum	514	29	65.9	685	8	AD011344	Ad011344	Novel	hum
442	29	65.9	685	7	AD015667	Ad015667	Novel	hum	515	29	65.9	685	8	AD013037	Ad013037	Human	PRO
443	29	65.9	685	7	AD016296	Ad016296	Novel	hum	516	29	65.9	685	8	AD068094	Ad068094	Novel	hum
444	29	65.9	685	7	AD015115	Ad015115	Novel	hum	517	29	65.9	685	8	AD015264	Ad015264	Human	PRO
445	29	65.9	685	7	AD014563	Ad014563	Novel	hum	518	29	65.9	685	8	AD012123	Ad012123	Novel	hum
446	29	65.9	685	7	AD048937	Ad048937	Novel	hum	519	29	65.9	685	8	AD094680	Ad094680	Novel	hum
447	29	65.9	685	8	AD0680825	Ad0680825	Novel	hum	520	29	65.9	685	8	AD056662	Ad056662	Human	PRO
448	29	65.9	685	8	AD021088	Ad021088	Novel	hum	521	29	65.9	685	8	AD050587	Ad050587	Human	PRO
449	29	65.9	685	8	AD050932	Ad050932	Human	PRO	522	29	65.9	685	8	AD023431	Ad023431	Novel	hum
450	29	65.9	685	8	AD076273	Ad076273	Human	PRO	523	29	65.9	685	8	AD063720	Ad063720	Novel	hum
451	29	65.9	685	8	AD075161	Ad075161	Human	PRO	524	29	65.9	685	8	AD024621	Ad024621	Novel	hum
452	29	65.9	685	8	AD075907	Ad075907	Novel	hum	525	29	65.9	685	8	AD094680	Ad094680	Novel	hum
453	29	65.9	685	8	AD085139	Ad085139	Novel	hum	526	29	65.9	685	8	AD060618	Ad060618	Novel	hum
454	29	65.9	685	8	AD086965	Ad086965	Novel	hum	527	29	65.9	685	8	AD007470	Ad007470	Novel	hum
455	29	65.9	685	8	AD020842	Ad020842	Novel	hum	528	29	65.9	685	8	AD060629	Ad060629	Novel	hum
456	29	65.9	685	8	AD039139	Ad039139	Novel	hum	529	29	65.9	685	8	AD061733	Ad061733	Novel	hum
457	29	65.9	685	8	AD087637	Ad087637	Human	PRO	530	29	65.9	685	8	AD081934	Ad081934	Human	PRO
458	29	65.9	685	8	AD086041	Ad086041	Human	PRO	531	29	65.9	685	8	AD057173	Ad057173	Novel	hum
459	29	65.9	685	8	AD056686	Ad056686	Human	PRO	532	29	65.9	685	8	AD056621	Ad056621	Novel	hum
460	29	65.9	685	8	AD073671	Ad073671	Human	PRO	533	29	65.9	685	8	AD055517	Ad055517	Novel	hum
461	29	65.9	685	8	AD075489	Ad075489	Human	PRO	534	29	65.9	685	8	AD058277	Ad058277	Novel	hum
462	29	65.9	685	8	AD078511	Ad078511	Novel	hum	535	29	65.9	685	8				

536	29	65.9	685	8	ADG70643	Adg70643	Novel hum	609	29	65.9	843	7	ADF90890	Adf90890	Human hep
537	29	65.9	685	8	ADH39120	Adh39120	Novel hum	610	29	65.9	843	8	ADN03920	Adn03920	Antipscori
538	29	65.9	685	8	ADG57725	Adg57725	Novel hum	611	29	65.9	843	8	ADZ87302	Adz87302	Human com
539	29	65.9	685	8	ADG53309	Adg53309	Novel hum	612	29	65.9	872	7	ADZ80879	Adz80879	Novel pro
540	29	65.9	685	8	ADG71195	Adg71195	Novel hum	613	29	65.9	1084	6	ABO01391	Ab001391	Human pro
541	29	65.9	685	8	ADG61382	Adg61382	Human PRO	614	29	65.9	1084	6	ADN96154	Adn96154	Human NOV
542	29	65.9	685	8	ADH30344	Adh30344	Human PRO	615	29	65.9	1086	4	ABE71547	Ab071547	Drosophill
543	29	65.9	685	8	ADH11711	Adh11711	Novel hum	616	29	65.9	1088	6	ABO01397	Ab001397	Human pro
544	29	65.9	685	8	ADG52133	Adg52133	Novel hum	617	29	65.9	1088	6	ADN96166	Adn96166	Human NOV
545	29	65.9	685	8	ADG53861	Adg53861	Novel hum	618	29	65.9	1401	4	ABB61146	Abb61146	Drosophill
546	29	65.9	685	8	ADG80830	Adg80830	Human PRO	619	29	65.9	1423	8	ADO44499	Ad044499	Murine sl
547	29	65.9	685	8	ADG56069	Adg56069	Novel hum	620	29	65.9	1564	7	ADC86801	Adc86801	Human GPC
548	29	65.9	685	8	ADH12335	Adh12335	Novel hum	621	29	65.9	1583	6	ABG76428	Abg76428	Apple etc
549	29	65.9	685	8	ADG61181	Adg61181	Novel hum	622	29	65.9	2743	5	ABB81598	Abb81598	Human lam
550	29	65.9	685	8	ADH28268	Adh28268	Human PRO	623	29	65.9	3597	5	ABO09503	Ab009503	Human lam
551	29	65.9	685	8	ADG54413	Adg54413	Novel hum	624	29	65.9	3597	5	ADO10047	Ad010047	Novel hum
552	29	65.9	685	8	ADG59453	Adg59453	Novel hum	625	29	65.9	3600	5	ABB09501	Abb09501	Novel hum
553	29	65.9	685	8	ADH433709	Adh433709	Human PRO	626	29	65.9	3600	8	ADO10043	Ad010043	Novel hum
554	29	65.9	685	8	ADG34210	Adg34210	Novel hum	627	29	65.9	3635	5	ABB81589	Abb81589	Mouse lam
555	29	65.9	685	8	ADH30877	Adh30877	Human PRO	628	29	65.9	3635	5	AAM50357	Aam50357	Mouse lam
556	29	65.9	685	8	ADH14423	Adh14423	Novel hum	629	29	65.9	3635	8	ADP74034	Adp74034	Murine la
557	29	65.9	685	8	ADH69774	Adh69774	Human PRO	630	29	65.9	3690	8	ADM87254	Adm87254	Human pro
558	29	65.9	685	8	ADG09620	Adg09620	Novel hum	631	29	65.9	3695	5	ABB81588	Abb81588	Human lam
559	29	65.9	685	8	ADH15091	Adh15091	Novel hum	632	29	65.9	3695	5	ADT28030	Adt28030	ECM/CAD pr
560	29	65.9	685	8	ADG08968	Adg08968	Novel hum	633	29	65.9	3695	5	ADN04478	Adn04478	Antipscori
561	29	65.9	685	8	ADH14423	Adh14423	Novel hum	634	29	65.9	3695	8	ADY25780	Ady25780	MRAC LAMA
562	29	65.9	685	8	ADH29935	Adh29935	Novel hum	635	29	65.9	3696	5	AAE17310	Aae17310	Human lam
563	29	65.9	685	8	ADH18018	Adh18018	Novel hum	636	29	65.9	3705	5	AAE17309	Aae17309	Human lam
564	29	65.9	685	8	ADM27332	Adm27332	Novel hum	637	29	65.9	4123	7	ABU62079	Abu62079	Human jel
565	29	65.9	685	8	ADH63289	Adh63289	Novel hum	638	29	65.9	4219	5	ADH48718	Adh48718	NOVI prot
566	29	65.9	685	8	ADH77194	Adh77194	Human PRO	639	29	65.9	4561	4	ABG30203	Abg30203	Novel hum
567	29	65.9	685	8	ADK83054	Adk83054	Human PRO	640	29	65.9	5737	7	ADN95328	Adn95328	Human BEC
568	29	65.9	685	8	ADK66690	Adk66690	Human PRO	641	29	65.9	9222	4	ABG21064	Abg21064	Novel hum
569	29	65.9	685	8	ADK00015	Adk00015	Human NOT	642	29	65.9	11	4	AAU04530	Aau04530	VEGF base
570	29	65.9	685	8	ADH65316	Adh65316	Human PRO	643	29	65.9	11	4	ABB45060	Abb45060	Rat album
571	29	65.9	685	8	ADM27452	Adm27452	Human PRO	644	29	65.9	12	8	ADJ73673	Adj73673	Rat serum
572	29	65.9	685	8	ADM41513	Adm41513	Human del	645	29	65.9	12	8	ADJ25802	Adj25802	Proline t
573	29	65.9	685	8	ADL73027	Adl73027	Human Not	646	29	65.9	14	9	ADY93363	Ady93363	Peptide 1
574	29	65.9	685	8	ADM42176	Adm42176	Human PRO	647	29	65.9	15	4	ABB45080	Abb45080	Rat album
575	29	65.9	685	8	ADM76013	Adm76013	Human Del	648	29	65.9	15	4	ABB45078	Abb45078	Rat album
576	29	65.9	685	8	ADM28038	Adm28038	Human PRO	649	29	65.9	15	8	ADJ73690	Adj73690	Rat serum
577	29	65.9	685	8	ADQ14406	Adq14406	Human Del	650	29	65.9	15	8	ADJ73692	Adj73692	Rat serum
578	29	65.9	685	8	ADR20594	Adr20594	Human Not	651	29	65.9	18	4	ABB45325	Abb45325	Human alb
579	29	65.9	685	8	ADR41796	Adr41796	Human Del	652	29	65.9	18	4	ABB45179	Abb45179	Rat album
580	29	65.9	685	8	ADI95520	Adi95520	Human PRO	653	29	65.9	18	4	ABB45197	Abb45197	Rat album
581	29	65.9	685	8	ADI96072	Adi96072	Novel hum	654	29	65.9	18	4	ABB45172	Abb45172	Rat album
582	29	65.9	685	8	ADR88260	Adr88260	Human Del	655	29	65.9	18	4	ABB45200	Abb45200	Rat album
583	29	65.9	685	8	ADR88276	Adr88276	Mus muscu	656	29	65.9	18	8	ADJ73804	Adj73804	Rat serum
584	29	65.9	685	8	ADR89382	Adr89382	Human Del	657	29	65.9	18	8	ADJ73768	Adj73768	Rat serum
585	29	65.9	685	8	ADH75458	Adh75458	Human Del	658	29	65.9	18	8	ADJ73775	Adj73775	Rat serum
586	29	65.9	685	8	ADS32024	Ads32024	Novel hum	659	29	65.9	18	8	ADJ73796	Adj73796	Rat serum
587	29	65.9	685	8	ADP03008	Adp03008	Human PRO	660	29	65.9	18	8	ADJ73802	Adj73802	Rat serum
588	29	65.9	685	9	ADX58107	Adx58107	Human PRO	661	29	65.9	18	8	ADJ73941	Adj73941	Human ser
589	29	65.9	685	9	ADH70419	Adh70419	Human Del	662	29	65.9	18	8	ADJ73793	Adj73793	Rat serum
590	29	65.9	685	9	ADZ03059	Adz03059	Human sec	663	29	65.9	20	4	ABB45310	Abb45310	Human alb
591	29	65.9	685	9	ABE13805	Abel3805	Cancer ce	664	29	65.9	20	4	ABB45152	Abb45152	Rat album
592	29	65.9	685	9	ABE17854	Abel7854	Human not	665	29	65.9	20	4	ABB45130	Abb45130	Rat album
593	29	65.9	686	3	AAV76911	Aav76911	Mouse not	666	29	65.9	20	8	ADJ73749	Adj73749	Rat serum
594	29	65.9	686	3	AAV68911	Aav68911	Amino aci	667	29	65.9	20	8	ADJ73727	Adj73727	Rat serum
595	29	65.9	686	6	ABU55881	Abu55881	Mouse not	668	29	65.9	25	8	ADJ73927	Adj73927	Human ser
596	29	65.9	686	7	ABW02798	Abw02798	Mouse del	669	29	65.9	25	8	ADG68099	Adg68099	Dog pancr
597	29	65.9	686	7	ABW02788	Abw02788	Mouse del	670	29	65.9	25	8	ADG68098	Adg68098	pig pancr
598	29	65.9	686	7	ABW02796	Abw02796	Mouse del	671	29	65.9	26	2	AAR95893	Aar95893	Oviduct c
599	29	65.9	686	7	ABW02795	Abw02795	Mouse del	672	29	65.9	33	4	ABB42483	Abb42483	Peptide #
600	29	65.9	686	7	ABW02797	Abw02797	Mouse del	673	29	65.9	33	4	ABM62296	Abm62296	Peptide #
601	29	65.9	686	7	ADH88261	Adh88261	Mus muscu	674	29	65.9	33	4	AAW76186	Aaw76186	Human bon
602	29	65.9	711	7	ABO61594	Ab061594	Human dia	675	29	65.9	33	4	AAW63372	Aaw63372	Human bra
603	29	65.9	741	7	ABO61594	Ab061594	Klebsiell	676	29	65.9	47	2	ABG57904	Abg57904	Human liv
604	29	65.9	758	3	ABM68685	Abm68685	Rice abio	677	29	65.9	48	7	AAV41497	Aav41497	Fragment
605	29	65.9	764	3	AAV68949	Aav68949	Fusion pr	678	29	65.9	48	7	ADZ08183	Adz08183	Novel pro
606	29	65.9	797	8	ABM84517	Abm84517	Human dia	679	29	65.9	54	4	AAW19876	Aaw19876	Peptide #
607	29	65.9	802	8	ABM84516	Abm84516	Human dia	680	29	65.9	54	4	ABB39808	Abb39808	Peptide #
608	29	65.9	803	4	ABB62003	Abb62003	Drosophill	681	29	65.9	54	4	ABB39808	Abb39808	Peptide #

682	28	63.6	54	4	AAM33398	Aam33398	Peptide #	755	28	63.6	122	3	AAg03985	Aag03985	Human sec
683	28	63.6	54	4	ABb24422	Abb24422	Protein #	756	28	63.6	122	4	ABb52644	Abb52644	Escherich
684	28	63.6	54	4	AAM73188	Aam73188	Human bon	757	28	63.6	125	8	ADu17067	Adu17067	Mb. tuberc
685	28	63.6	54	4	AAU46504	Aau46504	Protonib	758	28	63.6	127	4	AAU42623	Aau42623	Protonib
686	28	63.6	54	4	AAM60532	Aam60532	Human bra	759	28	63.6	127	6	AMh39142	Amh39142	Protonib
687	28	63.6	54	4	ABG54904	Abg54904	Human liv	760	28	63.6	138	4	AAU29758	Aau29758	Novel hum
688	28	63.6	54	5	ABG43035	Abg43035	Human pep	761	28	63.6	138	9	AEA20886	Aea20886	Novel hum
689	28	63.6	54	6	ABM43123	Abm43123	Protonib	762	28	63.6	139	4	AAO10791	Aao10791	Human pol
690	28	63.6	56	6	ABP99771	Abp99771	Human sec	763	28	63.6	142	7	ABO74248	AbO74248	Protonib
691	28	63.6	56	6	ABR01263	AbR01263	Human gen	764	28	63.6	144	7	ABO81545	AbO81545	Pseudomon
692	28	63.6	56	7	ADC20517	AdC20517	Human sec	765	28	63.6	144	8	ADP31453	Adp31453	Human sec
693	28	63.6	57	3	AAb30084	Aab30084	Human sec	766	28	63.6	149	4	AAE04263	Aae04263	Human gen
694	28	63.6	58	6	ABP77249	Abp77249	N. gonorr	767	28	63.6	154	2	AAy33504	Aay33504	Human unl
695	28	63.6	63	3	AAb51734	Aab51734	Human sec	768	28	63.6	156	7	ABO74152	AbO74152	Pseudomon
696	28	63.6	63	3	AAU58881	Aau58881	Protonib	769	28	63.6	159	4	ABG13130	AbG13130	Novel hum
697	28	63.6	63	6	ABM55400	Abm55400	Protonib	770	28	63.6	164	4	AAE04265	Aae04265	Human gen
698	28	63.6	64	4	ABB16042	Abb16042	Human ner	771	28	63.6	164	4	AAE01650	Aae01650	Human gen
699	28	63.6	64	4	AAU57824	Aau57824	Protonib	772	28	63.6	166	8	ADX78647	Adx78647	Plant ful
700	28	63.6	69	6	ABM54343	Abm54343	Protonib	773	28	63.6	166	8	ADX71945	Adx71945	Plant ful
701	28	63.6	71	4	AAU52868	Aau52868	Protonib	774	28	63.6	171	8	ADP30794	Adp30794	Human sec
702	28	63.6	71	6	ABM49387	Abm49387	Protonib	775	28	63.6	171	8	ADP30793	Adp30793	Human sec
703	28	63.6	74	4	ABB66555	Abb66555	Drosophil	776	28	63.6	172	4	AAU39136	Aau39136	Human sec
704	28	63.6	74	4	AAU56073	Aau56073	Protonib	777	28	63.6	172	6	ABM35655	Abm35655	Protonib
705	28	63.6	74	6	ABM52592	Abm52592	Protonib	778	28	63.6	177	2	AAy36634	Aay36634	Human nov
706	28	63.6	74	6	ABP56104	Abp56104	Human IGF	779	28	63.6	177	6	ADA11875	Ada11875	Human nov
707	28	63.6	76	6	AAU42937	Aau42937	Protonib	780	28	63.6	180	7	ABR62395	AbR62395	Polypepti
708	28	63.6	76	6	ABM39456	Abm39456	Protonib	781	28	63.6	180	8	ADJ58667	Adj58667	Human NOV
709	28	63.6	76	8	ABO56663	AbO56663	Human gen	782	28	63.6	180	8	ADM90897	Adm90897	Human pha
710	28	63.6	79	4	AAU45268	Aau45268	Protonib	783	28	63.6	183	8	ADR66909	Adr66909	Human pro
711	28	63.6	79	6	ABM41787	Abm41787	Protonib	784	28	63.6	183	8	ADR66011	Adr66011	Human pro
712	28	63.6	80	6	AAU49012	Aau49012	Protonib	785	28	63.6	189	4	AAE04268	Aae04268	Human gen
713	28	63.6	80	6	ABM45531	Abm45531	Protonib	786	28	63.6	190	4	AAE01680	Aae01680	Human gen
714	28	63.6	83	4	AAU58896	Aau58896	Protonib	787	28	63.6	190	5	ABG63945	AbG63945	Human alb
715	28	63.6	83	6	ABM55415	Abm55415	Protonib	788	28	63.6	190	8	ADL77210	AdL77210	Albumin f
716	28	63.6	84	4	AAW96592	Aaw96592	Human rep	789	28	63.6	190	9	ADY18898	Ady18898	PRO polyp
717	28	63.6	84	7	ABO84011	AbO84011	Pseudomon	790	28	63.6	191	4	AAE04266	Aae04266	Human gen
718	28	63.6	86	3	AAAB4184	AaB4184	Human can	791	28	63.6	192	8	ADY12911	Ady12911	Plant ful
719	28	63.6	87	4	AAU67356	Aau67356	Protonib	792	28	63.6	199	2	AAE60316	Aae60316	Sheep LFA
720	28	63.6	87	6	ABM63875	Abm63875	Protonib	793	28	63.6	199	8	ADU17821	AdU17821	Reverse t
721	28	63.6	89	4	AAE04225	Aae04225	Human gen	794	28	63.6	205	2	AAy41495	Aay41495	Fragment
722	28	63.6	89	4	ABB38046	Abb38046	Peptide #	795	28	63.6	206	8	ADQ20449	AdQ20449	Human PRO
723	28	63.6	89	4	AAW31466	Aaw31466	Peptide #	796	28	63.6	207	2	AAW55459	Aaw55459	H. pylori
724	28	63.6	89	4	AAW71185	Aaw71185	Human bon	797	28	63.6	207	4	AAW83836	Aaw83836	Human lmm
725	28	63.6	89	4	AAW58678	Aaw58678	Human bra	798	28	63.6	218	2	AAW55202	Aaw55202	H. pylori
726	28	63.6	89	4	ABG52887	Abg52887	Human liv	799	28	63.6	218	8	AAy65681	Aay65681	C. elegans
727	28	63.6	89	5	ABG40982	Abg40982	Human pep	800	28	63.6	228	8	ADP30764	Adp30764	Human sec
728	28	63.6	89	5	ABG64494	Abg64494	Human alb	801	28	63.6	228	8	ADX69140	Adx69140	Plant ful
729	28	63.6	89	8	ADL77761	AdL77761	Albumin f	802	28	63.6	233	3	AAE09755	Aae09755	Human sec
730	28	63.6	91	4	AAU64239	Aau64239	Protonib	803	28	63.6	233	3	AAy67291	Aay67291	Human sec
731	28	63.6	91	6	ABM60758	Abm60758	Protonib	804	28	63.6	237	3	AAAB09620	AaB09620	Insulin-1
732	28	63.6	95	4	ADG27779	Adg27779	Human nov	805	28	63.6	237	4	ABG00543	AbG00543	Novel hum
733	28	63.6	100	4	AAE04262	Aae04262	Human gen	806	28	63.6	237	8	ADP30792	Adp30792	Human sec
734	28	63.6	104	3	AAAB2798	AaB2798	Human sec	807	28	63.6	237	9	ABW93777	AbW93777	Human dia
735	28	63.6	108	2	AAW07693	Aaw07693	Dendroide	808	28	63.6	237	9	ABW93777	AbW93777	Human dia
736	28	63.6	108	7	ADM05568	Adm05568	Human pro	809	28	63.6	238	8	ABW84306	AbW84306	Human dia
737	28	63.6	109	2	AAW07694	Aaw07694	Dendroide	810	28	63.6	248	7	ADCO1611	AdCO1611	Enterohae
738	28	63.6	109	8	ADP20690	Adp20690	Human sec	811	28	63.6	248	7	ADCO1611	AdCO1611	Enterohae
739	28	63.6	110	4	ABB65848	Abb65848	Drosophil	812	28	63.6	248	8	ADP30662	Adp30662	Human sec
740	28	63.6	110	4	ABBI1699	Abbi1699	Human GAB	813	28	63.6	248	8	ADP31443	Adp31443	Human sec
741	28	63.6	110	6	ABJ19814	Abj19814	Androgen-	814	28	63.6	250	8	ADM90898	Adm90898	Human pha
742	28	63.6	111	4	AAW14931	Aaw14931	Peptide #	815	28	63.6	254	4	AAy67290	Aay67290	Rat insul
743	28	63.6	111	4	ABBI9339	Abbi9339	Protein #	816	28	63.6	255	4	AAy67292	Aay67292	Human col
744	28	63.6	111	4	AAW67069	Aaw67069	Human bon	817	28	63.6	257	8	ADO70024	AdO70024	Human ins
745	28	63.6	111	4	AAW54666	Aaw54666	Human bra	818	28	63.6	258	2	AAE22253	Aae22253	Sequence
746	28	63.6	111	4	ABG48733	Abg48733	Human liv	819	28	63.6	258	2	AAE21688	Aae21688	Sequence
747	28	63.6	111	5	AAW02659	Aaw02659	Peptide #	820	28	63.6	258	2	AAW37466	Aaw37466	Inhibitor
748	28	63.6	113	3	ABG36727	Abg36727	Human pep	821	28	63.6	258	2	AAy33968	Aay33968	A human i
749	28	63.6	116	4	ABG19371	Abg19371	Novel hum	822	28	63.6	258	4	AAy67292	Aay67292	Insulin-1
750	28	63.6	116	4	ADP30955	Adp30955	Human sec	823	28	63.6	258	4	AAy67292	Aay67292	Insulin-1
751	28	63.6	117	8	AAU61137	Aau61137	Protonib	824	28	63.6	258	4	AAy67292	Aay67292	Insulin-1
752	28	63.6	117	6	ABM57656	Abm57656	Protonib	825	28	63.6	258	4	AAy67292	Aay67292	Insulin-1
753	28	63.6	117	6	AAO16319	Aao16319	Human ins	826	28	63.6	258	4	AAy67292	Aay67292	Insulin-1
754	28	63.6	117	6	AAO16319	Aao16319	Human ins	827	28	63.6	258	5	AAU86154	Aau86154	Human PRO

828	28	63.6	258	6	ABU71442	Abu71442 Human neo	901.	28	63.6	410	4	AAE00849	Aae00849 Human nov
829	28	63.6	258	7	ABR62394	Abri62394 Polypepti	902	28	63.6	423	8	ADP30819	Adp30819 Human sec
830	28	63.6	258	7	ADJ37335	Adj37335 Human tum	903	28	63.6	429	8	ADP30570	Adp30570 Human sec
831	28	63.6	258	8	ADH31398	Adh31398 Human mal	904	28	63.6	437	2	AAW67722	Aaw67722 Human tum
832	28	63.6	258	8	ADG68259	Adg68259 Human PRO	905	28	63.6	437	3	AAI87750	Aai87750 Human PRO
833	28	63.6	258	8	ADJ58665	Adj58665 Human NOV	906	28	63.6	437	3	AAI66737	Aai66737 Membrane-
834	28	63.6	258	8	ADP30597	Adp30597 Human sec	907	28	63.6	437	3	AAI96733	Aai96733 PRO1863,
835	28	63.6	258	8	ADY14554	Ady14554 PRO polyP	908	28	63.6	437	3	AAI87261	Aai87261 Human sig
836	28	63.6	258	9	ADY19576	Ady19576 PRO polyP	909	28	63.6	437	3	AAU12404	Aau12404 Human PRO
837	28	63.6	258	9	ADZ09729	Adz09729 Human bre	910	28	63.6	437	4	AAI70851	Aai70851 Human PRO
838	28	63.6	258	9	AEA15085	Aea15085 Human pol	911	28	63.6	437	4	AAE65260	Aae65260 Human PRO
839	28	63.6	260	8	ADK76813	Adk76813 Plant ful	912	28	63.6	437	5	AAH80767	Aah80767 Human gita
840	28	63.6	261	8	ADK67963	Adk67963 Plant ful	913	28	63.6	437	5	AAU083672	Aau083672 Human PRO
841	28	63.6	263	2	AAE60474	Aae60474 E. plimnat	914	28	63.6	437	5	ABG92707	Abg92707 Human sec
842	28	63.6	265	2	AEA19937	Aea19937 Novel hum	915	28	63.6	437	5	ABG91359	Abg91359 Novel hum
843	28	63.6	268	4	ABE66405	Abbe66405 Drosophi	916	28	63.6	437	5	ABG31399	Abg31399 Human PRO
844	28	63.6	270	4	ABG28002	Abg28002 Novel hum	917	28	63.6	437	5	ADJ31302	Adj31302 Human PRO
845	28	63.6	271	7	ADK42690	Adk42690 S. xenoph	918	28	63.6	437	6	ABU72375	Abu72375 Novel hum
846	28	63.6	273	2	AAI50557	Aai50557 H6V1 IAT	919	28	63.6	437	6	ABU58075	Abu58075 Human PRO
847	28	63.6	276	8	ADK73340	Adk73340 Plant ful	920	28	63.6	437	6	ABU59153	Abu59153 Novel hum
848	28	63.6	277	8	ADP30503	Adp30503 Human sec	921	28	63.6	437	6	ABU80865	Abu80865 Human sec
849	28	63.6	277	8	ADK67533	Adk67533 Plant ful	922	28	63.6	437	6	ABU82665	Abu82665 Human sec
850	28	63.6	279	8	ADP31509	Adp31509 Human sec	923	28	63.6	437	6	ABO17848	AbO17848 Novel hum
851	28	63.6	279	8	ADP31523	Adp31523 Human sec	924	28	63.6	437	6	ABU60584	Abu60584 Human sec
852	28	63.6	279	8	ADP31503	Adp31503 Human sec	925	28	63.6	437	6	ABU80819	Abu80819 Human PRO
853	28	63.6	280	4	ABG04786	Abg04786 Novel hum	926	28	63.6	437	6	ABO33785	AbO33785 Novel hum
854	28	63.6	280	4	ADK77420	Adk77420 Plant ful	927	28	63.6	437	6	ABU13966	Abu13966 Human PRO
855	28	63.6	282	7	ADK77654	Adk77654 Plant ful	928	28	63.6	437	6	ABU81102	Abu81102 Human PRO
856	28	63.6	295	7	ABO69521	AbO69521 Pseudomon	929	28	63.6	437	6	ABU72551	Abu72551 Novel hum
857	28	63.6	296	6	ABP97797	Abp97797 Amilo act	930	28	63.6	437	6	ABU66802	Abu66802 Human PRO
858	28	63.6	297	3	AAI33201	Aai33201 Zea mays	931	28	63.6	437	6	ABG73312	Abg73312 Human PRO
859	28	63.6	300	8	ADP30805	Adp30805 Human sec	932	28	63.6	437	6	ABU59883	Abu59883 Novel sec
860	28	63.6	302	7	ABR61804	Abri61804 Human JAG	933	28	63.6	437	6	ABU59300	Abu59300 Human sec
861	28	63.6	302	7	ABR61734	Abri61734 Human JAG	934	28	63.6	437	6	ABO25997	AbO25997 Human PRO
862	28	63.6	303	4	ABG19635	Abg19635 Novel hum	935	28	63.6	437	6	ABO25073	AbO25073 Human sec
863	28	63.6	304	8	ADP31659	Adp31659 Human sec	936	28	63.6	437	6	ABU82128	Abu82128 Novel hum
864	28	63.6	306	8	ADP31205	Adp31205 Human sec	937	28	63.6	437	6	ABU59006	Abu59006 Human sec
865	28	63.6	309	6	ABU27667	Abu27667 Protein e	938	28	63.6	437	6	ABU60811	Abu60811 Human sec
866	28	63.6	312	8	ADP31454	Adp31454 Human sec	939	28	63.6	437	6	ABU92384	Abu92384 Novel hum
867	28	63.6	315	8	ADP31685	Adp31685 Human sec	940	28	63.6	437	6	ABU59449	Abu59449 Novel hum
868	28	63.6	326	6	ABG99903	Abg99903 Human nov	941	28	63.6	437	6	ABU67078	Abu67078 Human sec
869	28	63.6	327	8	ADP30478	Adp30478 Human sec	942	28	63.6	437	6	ABU81234	Abu81234 Human PRO
870	28	63.6	328	6	ADB09578	Adb09578 Human sec	943	28	63.6	437	6	ABU92215	Abu92215 Novel hum
871	28	63.6	331	6	ADB09580	Adb09580 Alloiococ	944	28	63.6	437	6	ABU10921	Abu10921 Human PRO
872	28	63.6	331	7	ABO70490	AbO70490 Pseudomon	945	28	63.6	437	6	ABU81673	Abu81673 Novel hum
873	28	63.6	337	3	AAH07589	Aah07589 Protein e	946	28	63.6	437	6	ABU86512	Abu86512 Human sec
874	28	63.6	338	7	ABO63655	AbO63655 Klebsiell	947	28	63.6	437	6	ABO34126	AbO34126 Human PRO
875	28	63.6	342	1	AAI91375	Aai91375 Vaccinia	948	28	63.6	437	6	ADA45985	Ada45985 Novel hum
876	28	63.6	348	8	ADP31374	Adp31374 Human sec	949	28	63.6	437	6	ADA76416	Ada76416 Human PRO
877	28	63.6	349	5	ABH08605	Abh08605 Hlrudin P	950	28	63.6	437	6	ABU72308	Abu72308 Human PRO
878	28	63.6	350	8	ADP31672	Adp31672 Human sec	951	28	63.6	437	6	ADA19066	Ada19066 Human PRO
879	28	63.6	354	6	ADA57025	Ada57025 Human sec	952	28	63.6	437	6	ADA61689	Ada61689 Homo sapi
880	28	63.6	354	6	ADK40876	Adk40876 Human sec	953	28	63.6	437	6	ADBI9474	Adbi9474 Novel hum
881	28	63.6	354	6	ABR47807	Abri47807 Human sec	954	28	63.6	437	6	ADK476416	Adk476416 Human PRO
882	28	63.6	354	7	ADK74192	Adk74192 Human sec	955	28	63.6	437	6	ADA86494	Ada86494 Novel hum
883	28	63.6	355	2	AAV36267	Aav36267 Human sec	956	28	63.6	437	6	ADBI16058	Adbi16058 Human PRO
884	28	63.6	355	6	ADA11615	Ada11615 Human nov	957	28	63.6	437	6	ADA337866	Ada337866 Human sec
885	28	63.6	357	8	ADP31223	Adp31223 Human sec	958	28	63.6	437	6	ADA47844	Ada47844 Human PRO
886	28	63.6	357	8	ADP30505	Adp30505 Human sec	959	28	63.6	437	6	ADA21552	Ada21552 Human sec
887	28	63.6	360	5	ABP69663	Abp69663 Human sec	960	28	63.6	437	6	ADA10339	Ada10339 Human sec
888	28	63.6	360	8	ADN04946	Adn04946 Antipsoi	961	28	63.6	437	6	ADA67639	Ada67639 Human PRO
889	28	63.6	360	8	ADK66522	Adk66522 Plant ful	962	28	63.6	437	6	ADB30646	Adb30646 Human PRO
890	28	63.6	363	4	ABG07807	Abg07807 Novel hum	963	28	63.6	437	6	ADA85942	Ada85942 Novel hum
891	28	63.6	366	8	ADP31139	Adp31139 Human sec	964	28	63.6	437	6	ADA17883	Ada17883 Human PRO
892	28	63.6	371	8	ADP31612	Adp31612 Human sec	965	28	63.6	437	6	ADA97154	Ada97154 Human PRO
893	28	63.6	378	8	ADP31375	Adp31375 Human sec	966	28	63.6	437	6	ADA79458	Ada79458 Human PRO
894	28	63.6	384	8	ADP31575	Adp31575 Human sec	967	28	63.6	437	6	ADA87597	Ada87597 Novel hum
895	28	63.6	385	8	ADT57543	Adt57543 Plant pol	968	28	63.6	437	6	ADBI16799	Adbi16799 Human PRO
896	28	63.6	387	7	ADC67455	Adc67455 Human GPC	969	28	63.6	437	6	ADA27991	Ada27991 Human sec
897	28	63.6	387	7	ADP31308	Adp31308 Human sec	970	28	63.6	437	6	ADA91891	Ada91891 Novel hum
898	28	63.6	397	4	AAE00850	Aae00850 Human nov	971	28	63.6	437	6	ADBI14954	Adbi14954 Human PRO
899	28	63.6	405	8	ADP31534	Adp31534 Human sec	972	28	63.6	437	6	ADBI18915	Adbi18915 Novel hum
900	28	63.6	407	8	ADY08554	Ady08554 Plant ful	973	28	63.6	437	6	ADA94130	Ada94130 Human PRO

974	28	63.6	437	6	ADB20026	Novel hum
975	28	63.6	437	6	ADB13338	Human PRO
976	28	63.6	437	6	ABO43381	Novel hum
977	28	63.6	437	6	ADA94571	Human sec
978	28	63.6	437	6	ADA74592	Human PRO
979	28	63.6	437	6	ADB24825	Human PRO
980	28	63.6	437	6	ADA82349	Human PRO
981	28	63.6	437	6	ADA75312	Human PRO
982	28	63.6	437	6	ADA85330	Novel hum
983	28	63.6	437	6	ADA84838	Novel hum
984	28	63.6	437	6	ADB30094	Human PRO
985	28	63.6	437	6	ADA80622	Human PRO
986	28	63.6	437	6	ADA75864	Human PRO
987	28	63.6	437	6	ADA38796	Human sec
988	28	63.6	437	6	ADA47089	Human PRO
989	28	63.6	437	6	ADB25385	Human PRO
990	28	63.6	437	6	ADA93561	Human PRO
991	28	63.6	437	6	ADB26911	Human PRO
992	28	63.6	437	6	ADB31198	Human PRO
993	28	63.6	437	6	ABJ72436	Human PRO
994	28	63.6	437	6	ABU62955	Human PRO
995	28	63.6	437	6	ADA92917	Human sec
996	28	63.6	437	6	ADA61126	Homo sapi
997	28	63.6	437	6	ADB24273	Human PRO
998	28	63.6	437	6	ADA96602	Human PRO
999	28	63.6	437	6	ADA81174	Human PRO
1000	28	63.6	437	6	ADA96050	Human PRO

ALIGNMENTS

RESULT 1
 ADS23148 standard; protein; 192 AA.
 ID ADS23148;
 AC ADS23148;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #12181.
 XX
 DE Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX
 OS Bacteria.
 OS
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.

XX
 PS Claim 1; SEQ ID NO 12181; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for producing plants with
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 192 AA;

Query Match 81.8%; Score 36; DB 8; Length 192;
 Best Local Similarity 71.4%; Pred. No. 91;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGCC 10
 Db 176 CSPATGCC 182

RESULT 2
 AAU49413
 ID AAU49413 standard; protein; 79 AA.
 XX
 AC AAU49413;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #10309.
 XX
 DE SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 OS
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US012865.
 XX
 PR 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59545.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.

XX Example 1; SEQ ID NO 10608; 10699p; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 79 AA;

Query Match 79.5%; Score 35; DB 4; Length 79;
Best Local Similarity 71.4%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
| | | | |
DB 3 CSPGTGC 9

RESULT 3

ID ABM45932 standard; protein; 79 AA.

XX ABM45932;

DT 20-OCT-2003 (first entry)

DE Propionibacterium acnes predicted ORF-encoded polypeptide #10608.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;

KW immunostimulant; immune response; vaccine.

OS Propionibacterium acnes.

PN WO200303515-A1.

XX 24-APR-2003.

PD 11-OCT-2002; 2002WO-US032727.

PR 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes MC, Benson DR, Jones R, Carter D;

XX Barch B, Valiave-Douglas J;

XX WPI; 2003-381789/36.

DR N-PSDB; ACF64474.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

PT or for stimulating an immune response specific for a P. acnes protein.

XX Example 1; SEQ ID NO 10608; 1481p; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 79 AA;

Query Match 79.5%; Score 35; DB 6; Length 79;
Best Local Similarity 71.4%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
| | | | |
DB 3 CSPGTGC 9

RESULT 4

ID ABO74428 standard; protein; 180 AA.

XX ABO74428;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #6603.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

PN US6551795-B1.

XX 22-APR-2003.

PD 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

PI WPI; 2003-615309/58.

DR N-PSDB; ABD07999.

XX

PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX
PS Disclosure; SEQ ID NO 23174; 455bp; English.
XX
XX
CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-*P. aeruginosa* drugs, as templates for recombinant
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology. Sequences AB067826-
CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 180 AA;
XX
XX
Query Match 79.5%; Score 35; DB 7; Length 180;
Best Local Similarity 71.4%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 2;
XX
QY 4 CXPXTGC 10
DB 124 CRPATGC 130
XX
XX
RESULT 5
ADT04024
ID ADT04024 standard; protein; 371 AA.
XX
XX
AC ADT04024;
XX
DT 30-DEC-2004 (first entry)
XX
DE Human protein modification and maintenance molecule protein SeqID4.
XX
XX
KW protein modification and maintenance molecule; PMMW; cytostatic;
KW immunomodulator; expression; immune disorder; cancer; human.
XX
XX
OS Homo sapiens.
XX
XX
PN WO2004084806-A2.
XX
PD 07-OCT-2004.
XX
XX
PF 16-MAR-2004; 2004WO-US08006.
XX
XX
PR 21-MAR-2003; 2003US-0456864P.
PR 03-APR-2003; 2003US-0460512P.
PR 19-MAY-2003; 2003US-0472827P.
PR 30-MAY-2003; 2003US-0475072P.
PR 30-MAY-2003; 2003US-0475190P.
PR 24-JUN-2003; 2003US-0482575P.
PR 13-AUG-2003; 2003US-0495151P.
XX
XX
PA (INCY-) INCYTE CORP.
XX
XX
PI Baughn MR, Marquis JP, Kabie AE, Chawla NK, Emerling BM, Lee SY,
PI Hafalia AJA, Ramkumar J, Richardson TW, Wang JT, Nakamura IL,
PI Yang J, Jin P, Becha SD, Wilson AD;
XX
XX
DR WPI; 2004-710256/69.
DR N-PSDB; ADT04052.
XX
PT New human protein modification and maintenance molecules (PMMW)

PT polypeptide, useful in preparing a composition for treating a disease
PT associated with decreased expression or overexpression of PMMW e.g.,
PT cancer.
XX
XX
PS Claim 1; SEQ ID NO 4; 222bp; English.
XX
XX
CC This invention relates to novel human protein modification and
CC maintenance molecule (PMMW) proteins and the DNA sequences which encode
CC them. The invention may be useful for the production of compounds with a
CC cytostatic or immunomodulator activity. The proteins are useful in
CC preparing a composition for diagnosing or treating a disease or condition
CC associated with decreased expression or overexpression of functional
CC PMMW, for example immune disorders or cancer. The present sequence is
CC that of a human protein modification and maintenance molecule (PMMW)
CC protein of the invention.
XX
SQ Sequence 371 AA;
XX
XX
Query Match 79.5%; Score 35; DB 8; Length 371;
Best Local Similarity 71.4%; Pred. No. 2.4e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 2;
XX
QY 4 CXPXTGC 10
DB 364 CSPHTGC 370
XX
XX
RESULT 6
AAV62764
ID AAV62764 standard; peptide; 7 AA.
XX
XX
AC AAV62764;
XX
XX
DT 02-MAR-2000 (first entry)
XX
XX
DE PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:4047.
XX
XX
KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW inhibition; cadherin extracellular domain; cell adhesion recognition;
KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW neurological disease; cyclic.
XX
XX
OS Synthetic.
XX
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Disulfide-bond 1..7
XX
XX
PN WO9957149-A2.
XX
XX
PD 11-NOV-1999.
XX
XX
PF 05-MAY-1999; 99WO-CA000363.
XX
XX
PR 05-MAY-1998; 98US-00073040.
PR 06-NOV-1998; 98US-00187859.
PR 20-JAN-1999; 99US-00234395.
PR 08-MAR-1999; 99US-00264516.
XX
XX
PA (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
XX
PI Blaschuk OW, Gour BJ, Byers S;
PI WPI; 2000-038791/03.
XX
XX
DR WPI; 2000-038791/03.
XX
PT New cadherin modulating agents, used for modulating nonclassical cadherin
PT mediated functions for treating e.g. cancers, obesity, rheumatoid
PT arthritis, multiple sclerosis, diabetes or a neurological disease.
XX

DT 05-SEP-2002 (first entry)
XX
DB B lymphocyte stimulator protein binding peptide #1.
XX
KW B lymphocyte stimulator protein binding protein; Blys; immune disease;
XX allergy; proliferative disease; infectious disease; arteriosclerosis;
KW inflammatory disorder; hypergammaglobulinaemia; blood clotting;
XX ischaemia; graft-versus-host disease; neurodegenerative disease;
OS immunosuppressive; nephrotropic; antineumatic; antiarthritic;
XX neuroprotective; cytostatic; immunostimulant; antitumour; anti-HIV;
KW antiasclerotic; antiallergic; thyromimetic; antianaemic; haemostatic;
KW dermatological; antinflammatory; cardiant; ophthalmological; uropathic;
XX antidiabetic; antithyroid; antidepressant; hepatotropic.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 2 /label= Phe, Trp, Tyr
FT Misc-difference 4 /label= Pro, Tyr
FT
FT
FT
FT
PN WO200216411-A2.
XX
XX 28-FEB-2002.
XX
XX 17-AUG-2001; 2001WO-US025850.
XX
XX 18-AUG-2000; 2000US-0226700P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Belzer JP, Potter DM, Fleming TL, Rosen CA;
XX
XX MPI; 2002-499775/53.
XX
XX The treatment of various diseases e.g. rheumatoid arthritis, comprises
XX administering B lymphocyte stimulator binding polypeptide.
XX
XX Claim 69; Page 233; 387pp; English.
XX
XX The present invention relates to the treatment, prevention or
XX amelioration of a disease or disorder associated with: aberrant B
XX lymphocyte stimulator (Blys), Blys receptor expression or activity; cells
XX of haematopoietic origin; or proliferative disease; and reducing,
XX inhibiting or stimulating immunoglobulin production, B cell proliferation
XX and graft rejection involving administration of Blys binding polypeptide.
XX The Blys binding polypeptides are used in the treatment, prevention or
XX amelioration of diseases such as immune system diseases, proliferative
XX diseases, diseases of cells of haematopoietic origin, graft rejection,
XX allergies, infectious diseases, arteriosclerosis, inflammatory disorders,
XX hypergammaglobulinaemia, blood clotting disorders, ischaemia, and
XX neurodegenerative diseases. The present sequence is a B lymphocyte
XX stimulator protein binding peptide
SQ
Sequence 7 AA:
Query Match 77.3%; Score 34; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 CXPXTGC 10
DB 1 CXPXTGC 7
RESULT 11
ABG33862 standard; peptide; 7 AA.
XX
AC ABG33862;
XX
DT 15-JUL-2002 (first entry)

XX
DE B lymphocyte stimulator (Blys) binding peptide #436.
XX
KW B lymphocyte stimulator protein; B lymphocyte stimulator binding peptide;
KW Blys; biological fluid; serum; plasma; lymph; blood; urine; spinal fluid;
KW synovial fluid; saliva; mucus.
XX
OS Synthetic.
XX
PN WO200216412-A2.
XX
XX 28-FEB-2002.
XX
XX 17-AUG-2001; 2001WO-US025891.
XX
XX 18-AUG-2000; 2000US-0226489P.
XX
XX (DYAX-) DYAX CORP.
XX
XX Belzer JP, Potter MD, Fleming TJ, Iadner RC;
XX
XX MPI; 2002-351647/38.
XX
XX New B lymphocyte stimulator binding polypeptide useful in detecting or
XX isolating Blys or Blys-like polypeptide comprises a specified amino acid
XX sequence.
XX
XX disclosure; Page 132; 263pp; English.
XX
XX The invention relates to a B lymphocyte stimulator (Blys) binding
XX polypeptide. Blys binding peptides bind Blys or Blys-like proteins
XX reversibly or irreversibly. The binding peptides are used in detection,
XX isolation and/or purification of Blys in a solution such as water or a
XX buffer solution, as well as any fluid and/or cell obtained from an
XX individual biological fluid, body tissue, body cell, cell line, tissue
XX culture or other source containing Blys or Blys-like polypeptides. The
XX biological fluids include sera, plasma, lymph, blood, blood fraction,
XX urine, synovial fluid, spinal fluid, saliva and mucous. Sequences
XX ABG33406-33415, ABG33423-33575, ABG33588-33846, ABG33848-33850 and
XX ABG33852-33862 represent Blys binding peptides of the invention
XX
SQ
Sequence 7 AA:
Query Match 77.3%; Score 34; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 CXPXTGC 10
DB 1 CXPXTGC 7
RESULT 12
ABJ00545 standard; peptide; 13 AA.
XX
AC ABJ00545;
XX
DT 05-SEP-2002 (first entry)
XX
DB B lymphocyte stimulator protein binding protein consensus peptide #1.
XX
XX B lymphocyte stimulator protein binding protein; Blys; immune disease;
XX allergy; proliferative disease; infectious disease; arteriosclerosis;
KW inflammatory disorder; hypergammaglobulinaemia; blood clotting;
KW ischaemia; graft-versus-host disease; neurodegenerative disease;
KW immunosuppressive; nephrotropic; antineumatic; antiarthritic;
KW neuroprotective; cytostatic; immunostimulant; antitumour; anti-HIV;
KW antiasclerotic; antiallergic; thyromimetic; antianaemic; haemostatic;
KW dermatological; antinflammatory; cardiant; ophthalmological; uropathic;
KW antidiabetic; antithyroid; antidepressant; hepatotropic.
XX
OS Unidentified.

```

XX Key Location/Qualifiers
FH Misc-difference 1 /label= Ala, Asn, Lys, Ser
FT Misc-difference 2 /label= Ala, Glu, Met, Ser, Val
FT Misc-difference 3 /label= Ala, Asn, Lys, Pro
FT Misc-difference 5 /label= Phe, Trp, Tyr
FT Misc-difference 7 /label= Pro, Tyr
FT Misc-difference 11 /label= Ala, Gln, His, Phe, Val
FT Misc-difference 12 /label= Asn, Gln, Gly, His, Ser, Val
FT Misc-difference 13 /label= Ala, Asn, Gly, Ile, Pro, Ser
FT WO200216411-A2.
PN 28-FEB-2002.
XX 17-AUG-2001; 2001WO-US025850.
XX 18-AUG-2000; 2000US-0226700P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Belzer JP, Potter DM, Fleming TL, Rosen CA;
XX WPI; 2002-499775/53.
XX The treatment of various diseases e.g. rheumatoid arthritis, comprises
PT administering B lymphocyte stimulator binding polypeptide.
XX Claim 69; Page 230; 387pp; English.
XX The present invention relates to the treatment, prevention or
CC amelioration of a disease or disorder associated with: aberrant B
CC lymphocyte stimulator (BLyS), BLyS receptor expression or activity; cells
CC of haematopoietic origin; or proliferative disease; and reducing,
CC inhibiting or stimulating immunoglobulin production, B cell proliferation
CC and graft rejection involving administration of BLyS binding polypeptide.
CC The BLyS binding polypeptides are used in the treatment, prevention or
CC amelioration of diseases such as immune system diseases, proliferative
CC diseases, diseases of cells of hematopoietic origin, graft rejection,
CC allergies, infectious diseases, arteriosclerosis, inflammatory disorders,
CC hypergammaglobulinaemia, blood clotting disorders, ischaemia, and
CC neurodegenerative diseases. The present sequence is a conserved region of
CC a B lymphocyte stimulator protein binding peptide
XX Sequence 13 AA;
SQ
Query Match 77.3%; Score 34; DB 5; Length 13;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 CXPXTGC 10
DB 4 CXPXTGC 10
RESULT 13
ID ABG33861 standard; peptide; 13 AA.
XX
XX ABG33861;
AC
XX 15-JUL-2002 (first entry)
XX
XX B lymphocyte stimulator (BLyS) binding peptide #435.
DE
XX

```

```

KW B lymphocyte stimulator protein; B lymphocyte stimulator binding peptide;
KW BLyS; biological fluid; serum; plasma; lymph; blood; urine; spinal fluid;
KW synovial fluid; saliva; mucus.
OS Synthetic.
XX WO200216412-A2.
XX
XX 28-FEB-2002.
XX
XX 17-AUG-2001; 2001WO-US025891.
XX
XX 18-AUG-2000; 2000US-0226489P.
XX
XX (DVAX-) DVAX CORP.
XX
XX Belzer JP, Potter MD, Fleming TL, Ladner RC;
XX WPI; 2002-351647/38.
XX
XX New B-lymphocyte stimulator binding polypeptide useful in detecting or
PT isolating BLyS or BLyS-like polypeptide comprises a specified amino acid
PT sequence.
XX Disclosure; Page 121; 269pp; English.
XX
XX The invention relates to a B lymphocyte stimulator (BLyS) binding
CC polypeptide. BLyS binding peptides bind BLyS or BLyS-like proteins
CC reversibly or irreversibly. The binding peptides are used in detection,
CC isolation and/or purification of BLyS in a solution such as water or a
CC buffer solution, as well as any fluid and/or cell obtained from an
CC individual biological fluid, body tissue, body cell, cell line, tissue
CC culture or other source containing BLyS or BLyS-like polypeptides. The
CC biological fluids include sera, plasma, lymph, blood, blood fraction,
CC urine, synovial fluid, spinal fluid, saliva and mucus. Sequences
CC ABG33406-33415, ABG33423-33575, ABG33588-33846, ABG33848-33850 and
CC ABG33852-33862 represent BLyS binding peptides of the invention
XX
XX Sequence 13 AA;
SQ
Query Match 77.3%; Score 34; DB 5; Length 13;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 CXPXTGC 10
DB 4 CXPXTGC 10
RESULT 14
ID AAU44932
XX AAU44932;
XX
XX 27-FEB-2002 (first entry)
XX
XX Propionibacterium acnes immunogenic protein #5828.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
XX
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US012865.
XX
XX 21-APR-2000; 2000US-0199047P.
XX

```


02-JUN-2000; 2000US-0208844P.
PR 07-JUL-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonmeuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
DR N-PSDB; AAS59524.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 6127; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 53 AA;

Query Match	77.3%	Score 34	DB 4	Length 53
Best Local Similarity	71.4%	Pred. No. 73		
Matches	5	Conservative	0	Mismatches 2
		Indels	0	Gaps 0

Qy	4	CXPXTGC	10
Db	33	CRPRTGC	39

RESULT 15
ABM41451
ID ABM41451 standard; protein; 53 AA.
XX
XX ABM41451;
XX
AC
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #6127.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
XX Propionibacterium acnes.
OS
PN WO2003033515-A1.
PD 24-APR-2003.
XX
XX
PF 11-OCT-2002; 2002WO-US023727.
XX
ER 15-OCT-2001; 2001US-00978925.
XX
XX (CORI-) CORIXA CORP.

XX Mitchem JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barch B, Vallilve-Douglas J;
 XX
 DR WPI: 2003-381789/36.
 DR N-PSDB; ACF6453.
 XX
 XX New Propionibacterium acnes polypeptidases and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX
 XX Example 1; SEQ ID NO 6127; 1481pp. English.
 XX

CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM55624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

	Query Match	77.3%	Score 34;	DB 6;	Length 53;
	Best local similarity	71.4%;	Pred. No. 73;		
Matches	5; Conservative	0;	Mismatches	2;	Indels 0; Gaps 0.
Oy	4 CXPYTC 10				
Db	33 CRPRTGC 39				
RESULT 16					
ID	ADX87568				
	ADX87568 standard; protein; 99 AA.				
XX	ADX87568;				
DT					
XX	21-APR-2005 (first entry)				
DE	Plant full length insert polypeptide seqid 50232.				
KM	plant protectant; plant growth regulant; gene therapy; plant;				
KM	recombinant DNA construct; physical array; plant breeding marker;				
KM	cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;				
KM	extreme osmotic condition; pathogen tolerance; pest tolerance;				
KM	growth rate; cell cycle pathway; disease resistance;				
KM	galactomannan production; lignin production; plant growth regulator;				
KM	yield; plant growth; plant development; seed oil; protein yield;				
KS	protein content.				
OS	unidentified.				

```

XX XX US2004034888-A1.
PN
XX
XX 19-FEB-2004.
PD
XX
XX 28-APR-2003; 2003US-00425114.
PF
XX 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
XX (LITUJ/) LITU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABAS/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
PI WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX Claim 1; SEQ ID NO 50232; 15pp; English.
PS
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
XX Sequence 99 AA;
SQ
XX
XX Query Match 77.3%; Score 34; DB 8; Length 99;
XX Best Local Similarity 71.4%; Pred. No. 1.2e+02;
XX Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 4 CXPXTGC 10
XX | | | |
Db 91 CXPXTGC 97
XX
XX RESULT 17
XX ADA54961
XX ADA54961 standard; protein; 113 AA.
XX
XX ADA54961;
XX
XX 20-NOV-2003 (first entry)
XX
XX Human protein, SEQ ID 2529.
XX
XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
XX Gene Therapy; human; secretory protein; membrane proteins; cancer;
XX inflammatory disease; osteoporosis; neurological disease.
XX
XX Homo sapiens.
XX

```

```

XX XX EP1293569-A2.
PN
XX
XX 19-MAR-2003.
PD
XX
XX 21-MAR-2002; 2002EP-00006586.
PF
XX 14-SEP-2001; 2001JP-00328381.
PR 24-JAN-2002; 2002US-0350435P.
XX
XX (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
PI Yamamoto Y, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
XX WPI; 2003-395539/38.
XX
XX N-PSDB; ADA53322.
XX
XX New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
XX Claim 14; SEQ ID NO 2529; 205pp; English.
PS
XX
XX The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
XX Sequence 113 AA;
SQ
XX
XX Query Match 77.3%; Score 34; DB 6; Length 113;
XX Best Local Similarity 71.4%; Pred. No. 1.4e+02;
XX Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 4 CXPXTGC 10
XX | | | |
Db 103 CXPXTGC 109
XX
XX RESULT 18
XX AAB42027
XX AAB42027 standard; protein; 138 AA.
XX
XX AAB42027;
XX
XX 08-FEB-2001 (first entry)
XX
XX Human ORF1791 polypeptide sequence SEQ ID NO:1582.
XX
XX Human; open reading frame; ORF; detection; cyostatic; hepatotropic;
XX vulnery; anti-parasitic; anti-parkinsonian; nootropic; neuroprotective;
XX anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiant;
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX hypotensive; dermatological; immunosuppressive; antiinflammatory;
XX antiviral; antibacterial; antifungal; antirheumatic; antihypoid;
XX antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX cholesterol ester storage; systemic lupus erythematosus; infection;
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX bone damage; cartilage damage; antiinflammatory disease; coagulation;
XX thrombosis; contraceptive.
XX
XX Homo sapiens.
XX
XX WO200058473-A2.
XX
XX 05-OCT-2000.
XX

```

XX 31-MAR-2000; 2000MO-US008621.
 PF 31-MAR-1999; 99US-0127607P.
 XX 02-APR-1999; 99US-0127636P.
 PR 05-APR-1999; 99US-0127728P.
 PR 30-MAR-2000; 2000US-00540763.
 XX (CURA-) CURAGEN CORP.
 XX
 PI Shinketsu RA, Leach M;
 XX WPI; 2000-602362/57.
 DR N-PSDB; AAC76236.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease.
 XX
 PS Claim 11; Page 2743; 5507P; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in ABA40237 to ABA43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antiparasitic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
 CC antiviral; antifungal; antineumatic; antihistoid; and antianemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
 XX
 SQ Sequence 138 AA;
 Query Match 77.3%; Score 34; DB 3; Length 138;
 Best Local Similarity 71.4%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 CXPXTCG 10
 DB 15 CGPRTGC 21
 RESULT 19
 ABO75329
 ID ABO75329 standard; protein; 250 AA.
 XX
 AC ABO75329;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Pseudomonas aeruginosa polypeptide #7504.
 XX
 KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN US6551795-B1.
 XX
 PD 22-APR-2003.
 XX
 PF 18-FEB-1999; 99US-00252991.
 XX

PR 18-FEB-1998; 98US-0074788P.
 PR 27-JUL-1998; 98US-0094190P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 XX WPI; 2003-615309/58.
 DR N-PSDB; ABD08900.
 XX
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX
 PS Disclosure; SEQ ID NO 24075; 455P; English.
 XX
 CC The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-
 CC ABO68396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX
 SQ Sequence 250 AA;
 Query Match 77.3%; Score 34; DB 7; Length 250;
 Best Local Similarity 71.4%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 CXPXTCG 10
 DB 27 CPPATGC 33
 RESULT 20
 ABB93198
 ID ABB93198 standard; protein; 312 AA.
 XX
 AC ABB93198;
 XX
 DT 31-MAY-2002 (first entry)
 XX
 DE Herbicidally active polypeptide SEQ ID NO 2409.
 XX
 KW Herbicidal; plant; agriculture; herbicide.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO200210210-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 28-AUG-2001; 2001MO-EP009892.
 XX
 PR 28-AUG-2001; 2001MO-EP009892.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Tiejien K, Weidner M;
 XX WPI; 2002-269010/31.
 DR
 XX
 PT Identifying plant target proteins for herbicidally active compounds,
 XX

PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
PS Claim 5; SEQ ID NO 2409; 261pp + Sequence Listing; English.
XX
XX
XX
CC The invention relates to identifying target proteins (AB90790-AB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
CC herbicides
XX
XX
SQ Sequence 312 AA;
Query Match 77.3%; Score 34; DB 5; Length 312;
Best Local Similarity 71.4%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 4 CXPXTGC 10
DB 208 CDPNTGC 214
RESULT 21
ID ADN17632 standard; protein; 532 AA.
XX ADN17632;
XX
XX ADN17632;
XX
XX 02-DEC-2004 (first entry)
XX
XX Bacterial polypeptide #285.
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polypeptide.
XX
XX Bacteria.
XX
XX OS
XX PN US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
XX MPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 285; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX
SQ Sequence 532 AA;
Query Match 77.3%; Score 34; DB 8; Length 532;
Best Local Similarity 71.4%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 4 CXPXTGC 10
DB 122 CHPVTGC 128
RESULT 22
ID ADF05702 standard; protein; 540 AA.
XX ADF05702;
XX
XX 12-FEB-2004 (first entry)
XX
XX Bacterial polypeptide #1815.
XX
XX DE
XX Proteus mirabilis infection; bacterial infection; antibacterial;
XX immunostimulant.
XX
XX KM
XX OS Proteus mirabilis.
XX
XX PN US6605709-B1.
XX
XX 12-AUG-2003.
XX
XX 05-APR-2000; 2000US-00543681.
XX
XX 09-APR-1999; 99US-0128706P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Breton GL;
XX
XX MPI; 2003-895291/82.
XX N-PSDB; ADF01530.
XX
XX New Proteus mirabilis polypeptides and polynucleotides, useful as
XX reagents for diagnosis of bacterial disease, as components of
XX antibacterial vaccines, as targets for antibacterial drugs, or as
XX biocontrol agents for plants.
XX
XX Disclosure; SEQ ID NO 5987; 870pp; English.
XX
XX The invention relates to new Proteus mirabilis polypeptides and
XX polynucleotides. The invention also relates to antibodies against the

CC polypeptides, methods for producing the polypeptides, a method of
CC generating vaccines for immunising an individual against *P. mirabilis*, a
CC method for evaluating a compound for the ability to bind a *P. mirabilis*
CC polypeptide and a method for screening test compounds for anti-bacterial
CC activity. The polypeptides and polynucleotides are useful as molecular
CC targets for diagnosing, preventing and treating pathological conditions
CC resulting from bacterial infection, as reagents for diagnosis of
CC bacterial diseases, as components of antibacterial vaccines, as targets
CC for antibacterial drugs or as bio-control agents for plants. This
CC sequence represents a *Proteus mirabilis* polypeptide of the invention.

XX SQ Sequence 540 AA;

Query Match 77.3%; Score 34; DB 7; Length 540;
Best Local Similarity 71.4%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
| | | | |
DB 130 CXPVTGC 136

RESULT 23

ABG28408
ID ABG28408 standard; protein; 3226 AA.

XX AC ABG28408;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #28399.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSB-) HYSBQ INC.

XX PI Drmanac RT, Liu C, Tang YF;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS92595.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.

XX PS Claim 20; SEQ ID NO 58767; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 3226 AA;

Query Match 77.3%; Score 34; DB 4; Length 3226;
Best Local Similarity 71.4%; Pred. No. 2.1e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
| | | | |
DB 1707 CXPVTGC 1713

RESULT 24

AAU40006
ID AAU40006 standard; protein; 158 AA.

XX AC AAU40006;

XX DT 13-FEB-2002 (first entry)

XX DE Propionibacterium acnes immunogenic protein #902.

XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX OS Propionibacterium acnes.

XX PN WO200181581-A2.

XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-US012865.

XX PR 21-APR-2000; 2000US-0199047P.

XX PR 02-JUN-2000; 2000US-0208841P.

XX PR 07-JUL-2000; 2000US-0216747P.

XX PA (CORI-) CORIXA CORP.

XX PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX DR WPI; 2001-616774/71.

XX DR N-PSDB; AAS59510.

XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris.

XX PS Example 1; SEQ ID NO 1201; 1069pp; English.

CC Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies

CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomannan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant protein
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC segdata.uspto.gov/sequence.html?docID=20040216190.

XX Sequence 239 AA:

Query Match 75.0%; Score 33; DB 8; Length 239;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
| | | | |
Db 141 CKPPTGC 147

RESULT 27

AAAG43626 standard; protein; 751 AA.

XX AAAG43626;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 54550.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0127462P.

XX 01-APR-1999; 99US-0128234P.

XX 06-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132865P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139493P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144614P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155559P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 04-OCT-1999; 99US-0157573P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158022P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159284P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.

PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160880P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 72.7%; Score 32; DB 3; Length 751;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 90 CSPRSGC 96
RESULT 28
AAC435625
ID AAC43625 standard; protein; 798 AA.
XX AAC43625;
AC 18-OCT-2000 (first entry)
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 54549.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140931P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145129P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158363P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.

PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160880P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161559P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 72.7%; Score 32; DB 3; Length 798;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTCG 10
| | | | |
| : | | |
Db 137 CSPTSCG 143

RESULT 29
ID AAG43624 standard; protein; 835 AA.
XX AAG43624;
XX
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 54548.
XX
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridization assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136783P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138647P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140893P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.

PR	27-JUL-1999	99US-0145918P
PR	27-JUL-1999	99US-0145918P
PR	28-JUL-1999	99US-0145951P
PR	02-AUG-1999	99US-0146386P
PR	02-AUG-1999	99US-0146386P
PR	02-AUG-1999	99US-0146386P
PR	03-AUG-1999	99US-0146389P
PR	04-AUG-1999	99US-0147038P
PR	04-AUG-1999	99US-0147204P
PR	04-AUG-1999	99US-0147302P
PR	05-AUG-1999	99US-0147192P
PR	05-AUG-1999	99US-0147260P
PR	06-AUG-1999	99US-0147303P
PR	06-AUG-1999	99US-0147461P
PR	09-AUG-1999	99US-0147493P
PR	09-AUG-1999	99US-0147935P
PR	10-AUG-1999	99US-0148317P
PR	11-AUG-1999	99US-0148317P
PR	12-AUG-1999	99US-0148311P
PR	13-AUG-1999	99US-0148665P
PR	13-AUG-1999	99US-0148664P
PR	16-AUG-1999	99US-0149368P
PR	17-AUG-1999	99US-0149175P
PR	18-AUG-1999	99US-0149426P
PR	20-AUG-1999	99US-0149722P
PR	20-AUG-1999	99US-0149723P
PR	20-AUG-1999	99US-0149929P
PR	23-AUG-1999	99US-0149932P
PR	23-AUG-1999	99US-0149930P
PR	25-AUG-1999	99US-0150566P
PR	26-AUG-1999	99US-0150884P
PR	27-AUG-1999	99US-0151065P
PR	27-AUG-1999	99US-0151066P
PR	27-AUG-1999	99US-0151080P
PR	30-AUG-1999	99US-0151030P
PR	31-AUG-1999	99US-0151438P
PR	01-SEP-1999	99US-0151930P
PR	07-SEP-1999	99US-0152363P
PR	10-SEP-1999	99US-0153070P
PR	11-SEP-1999	99US-0153787P
PR	15-SEP-1999	99US-0154018P
PR	15-SEP-1999	99US-0154039P
PR	20-SEP-1999	99US-0154779P
PR	22-SEP-1999	99US-0155139P
PR	23-SEP-1999	99US-0155486P
PR	24-SEP-1999	99US-0155569P
PR	28-SEP-1999	99US-0156458P
PR	29-SEP-1999	99US-0156596P
PR	04-OCT-1999	99US-0157117P
PR	05-OCT-1999	99US-0157753P
PR	06-OCT-1999	99US-0157865P
PR	07-OCT-1999	99US-0158029P
PR	08-OCT-1999	99US-0158332P
PR	12-OCT-1999	99US-0158369P
PR	13-OCT-1999	99US-0159293P
PR	13-OCT-1999	99US-0159294P
PR	13-OCT-1999	99US-0159295P
PR	18-OCT-1999	99US-0159584P
PR	21-OCT-1999	99US-0159741P
PR	21-OCT-1999	99US-0160767P
PR	21-OCT-1999	99US-0160768P
PR	21-OCT-1999	99US-0160770P
PR	21-OCT-1999	99US-0159637P
PR	21-OCT-1999	99US-0160814P
PR	22-OCT-1999	99US-0160815P
PR	22-OCT-1999	99US-0160980P
PR	22-OCT-1999	99US-0160981P
PR	22-OCT-1999	99US-0160989P
PR	25-OCT-1999	99US-0161404P
PR	25-OCT-1999	99US-0161405P

PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161932P.
PR	29-OCT-1999;	99US-0162142P.
Query Match		
Best Local Similarity	72.7%;	Score 32; DB 3; Length 835;
Matches	4; Conservative	1; Mismatches 2; Indels 0; Gaps 0
OY	4 CXPXTGC 10	
DB	174 CSPSTGC 180	
RESULT 30		
AAY63264		
ID	AAY63264 standard; peptide; 7 AA.	
AC	AAY63264;	
XX		
DT	02-MAR-2000 (first entry)	
DE	Protocadherin cell adhesion recognition cyclic peptide SEQ ID NO:2748.	
XX		
KW	Modulation; nonclassical cadherin mediated cell adhesion; CAR;	
KW	inhibition; cadherin extracellular domain; cell adhesion recognition;	
KW	OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;	
KW	cadherin-14; cadherin-15; T-cadherin; PB-cadherin;	
KW	cadherin related neuronal receptor; LI-cadherin; protocadherin;	
KW	desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;	
KW	rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;	
KX	neurological disease; cyclic.	
OS	Synthetic.	
OS	Homo sapiens.	
FH		
FT	Disulfide-bond	Location/Qualifiers
PN	WC0957149-A2.	
PD	11-NOV-1999.	
XX		
XX	05-MAY-1999;	99WO-CA000363.
XX		
PR	05-MAY-1998;	98US-00073040.
PR	06-NOV-1998;	98US-00187859.
PR	20-JAN-1999;	99US-00234395.
PR	08-MAR-1999;	99US-00264516.
PA	(ADHER-) ADHEREX TECHNOLOGIES INC.	
PI	Blaschuk OW, Gour BJ, Byers S;	
PI	WPI; 2000-038791/03.	
XX		
DR		
XX		
PT	New cadherin modulating agents, used for modulating nonclassical cadherin-	
PT	mediated functions for treating e.g. dancers, obesity, rheumatoid	
PT	arthritis, multiple sclerosis, diabetes or a neurological disease.	
XX		
PS	Claim 84; Page 201; 252pp; English.	
XX		
CC	The present invention describes cadherin modulating agents (MA)	
CC	comprising peptides which comprise a nonclassical cadherin cell adhesion	
CC	recognition (CAR) sequence. The MAs can be used for modulating	
CC	nonclassical cadherin-mediated functions. They can be used for e.g.	
CC	inhibiting adhesion of nonclassical-cadherin expressing cells in a	
CC	mammal, enhancing delivery of a drug through the skin of a mammal,	
CC	enhancing delivery of a drug to a tumour in a mammal, treating cancer in	

CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC expressing cell, preventing or treating obesity in a mammal, stimulating
CC blood vessel regression in a mammal, enhancing drug delivery to the
CC central nervous system, treating a demyelinating neurological disease,
CC increasing vasopermeability in a mammal, enhancing synaptic stability in
CC nonclassical cadherin-expressing cells, inhibiting pregnancy in a
CC mammal, or preventing pregnancy in a mammal. They can also be used for
CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC -related macular degeneration, multiple sclerosis and diabetes. The
CC products can also be used for detection and diagnosis and in bioreactors.
CC AA60592 to AA64572 represent specifically claimed peptides, and
CC AA64573 to AA64643 and AA633183 to AA633186 represent sequences used in
CC the exemplification of the present invention

SQ Sequence 7 AA;

Query Match 70.5%; Score 31; DB 3; Length 7;
Best Local Similarity 57.1%; Pred. No. 2e+06;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
| | | |
| : ||
Db 1 CDPSSGC 7

RESULT 31

ABO81251
ID ABO81251 standard; protein; 90 AA.

AC ABO81251;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #13426.

KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI; 2003-615309/58.

DR N-PSDB; ABD14822.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 29997; 455dp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target

CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84336 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

SQ Sequence 90 AA;

Query Match 70.5%; Score 31; DB 7; Length 90;
Best Local Similarity 57.1%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
| | | |
| : ||
Db 29 CTPSAGC 35

RESULT 32

ABP06368
ID ABP06368 standard; protein; 114 AA.

AC ABP06368;

DT 25-JUN-2002 (first entry)

DE Human ORFX protein sequence SEQ ID NO:12718.

KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; hemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myaesthesia gravis.

XX Homo sapiens.

XX WO200192523-A2.

XX 06-DEC-2001.

PF 29-MAY-2001; 2001WO-US010836.

PR 30-MAY-2000; 2000US-0206132P.

PR 29-AUG-2000; 2000US-0228716P.

XX (CURA-) CURAGEN CORP.

PI Shinkets RA, Leach MD;

DR WPI; 2002-106308/14.

DR N-PSDB; AEN22120.

PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.

XX Disclosure; SEQ ID NO 12718; 1037pp; English.

XX The present invention describes substantially purified human proteins
XX (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
XX in the specification). AEN15762 to AEN27252 encode the human ORFX
XX proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
XX treating or preventing a pathology associated with an ORFX-associated
XX disorder in humans, and in the manufacture of a medicament for treating a
XX syndrome associated with ORFX-associated disorder. ORFX polynucleotide
XX sequences can be used in gene therapy. ORFX sequences can be used in the
XX treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
XX psoriasis, benign tumours, keloid, degenerative disorders, hemorrhage,
XX osteoarthritis, neurodegenerative disorders, disorders related to organ

CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease, and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 114 AA;

Query Match 70.5%; Score 31; DB 5; Length 114;
 Best Local Similarity 57.1%; Pred. No. 4.7e+02;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 CKPXTGC 10
 | | | | |
 Db 83 CGPTSGC 89

RESULT 33

ID ADJ71040 standard; protein; 148 AA.

ADJ71040;

06-MAY-2004 (first entry)

Human heat mitochondrial protein as a therapeutic target SeqID2846.

mitochondrial; human; screening assay; diabetes mellitus;

Huntington's disease; osteoarthritis;

Leber's hereditary optic neuropathy; LHON;

mitochondrial encephalopathy lactic acidosis and stroke; MELAS;

myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;

neuroprotective; neurotropic; antidiabetic; anticonvulsant; antiarthritic;

osteopathic; ophthalmological; cyostatic.

Homo sapiens.

MO2003087768-A2.

23-OCT-2003.

04-APR-2003; 2003WO-US010870.

12-APR-2002; 2002US-0372843P.

17-JUN-2002; 2002US-038987P.

20-SEP-2002; 2002US-0412418P.

(MITO-) MITOKOR.

(BUCK-) BUCK INST AGE RES.

Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;

Warnock DE;

WPI: 2003-845369/78.

Identifying a mitochondrial target for drug screening assays and for
 treating diseases associated with altered mitochondrial function,
 comprises detecting a modified polypeptide in a sample and correlating
 with the disease.

Claim 1; SEQ ID NO 2846; 180pp; English.

This invention relates to novel mitochondrial targets that can be used
 for therapeutic intervention in treating a disease associated with
 altered mitochondrial function. Specifically, it refers to a method for

CC identifying proteins of the human heart mitochondrial proteome that are
 CC useful for drug screening assays, as well as therapeutic targets. The
 CC present invention describes a method for identifying such proteins that
 CC can be used in the treatment of various diseases associated with altered
 CC mitochondrial function including diabetes mellitus, Huntington's disease,
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 CC compositions have neuroprotective, neurotropic, antidiabetic,
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
 CC cyostatic activities. This polypeptide sequence is a human heart
 CC mitochondrial protein of the invention.

Sequence 148 AA;

Query Match 70.5%; Score 31; DB 7; Length 148;
 Best Local Similarity 57.1%; Pred. No. 5.9e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 CKPXTGC 10
 | | | | |
 Db 69 CAPMAGC 75

RESULT 34

ID ADF58925 standard; protein; 183 AA.

ADF58925;

12-FEB-2004 (first entry)

Human polypeptide sequence SEQ ID NO:1333.

biological activity; genetic engineering; hybridization probe; oligomer;

primer; chromosome mapping; gene mapping; recombinant protein production;

human.

Homo sapiens.

MO2003080795-A2.

02-OCT-2003.

09-AUG-2002; 2002WO-US025485.

09-AUG-2001; 2001US-0311261P.

(HYSE-) HYSEQ INC.

Tang YT, Yang Y, Wang Z, Weng G, Ma Y;

WPI: 2003-876918/81.

N-PSDB; ADF57925.

New polynucleotides, useful as hybridization probes, oligomers or
 primers, for chromosome or gene mapping, for the recombinant production
 of proteins, and for generating antisense DNA or RNA.

Claim 20; SEQ ID NO 1333; 571pp; English.

The present sequence represents a polypeptide (II) with biological
 activity, which is encoded by an isolated polynucleotide sequence (I)
 from the present invention. Also described: (1) a vector comprising (I);
 (2) an expression vector comprising (I); (3) a host cell genetically
 engineered to comprise (I) which is operatively associated with a
 regulatory sequence that modulates expression of (I) in the host cell;
 (4) a polypeptide (II) encoded by (I); (5) a composition comprising the
 polypeptide of (4) and a carrier; (6) an antibody directed against the
 polypeptide of (4); (7) detecting (I) or the polypeptide of (4) in a
 sample; (8) identifying a compound that binds to the polypeptide of (4);
 (9) producing the polypeptide of (4); and (10) a collection of
 polynucleotides comprising at least one of the polynucleotide sequences

CC (1). The polynucleotides (I) can be used as hybridisation probes,
 CC oligomers or primers, for chromosome or gene mapping, for the recombinant
 CC production of proteins, and for generating antisense DNA or RNA.

XX SQ Sequence 183 AA;

Query Match 70.5%; Score 31; DB 7; Length 183;
 Best Local Similarity 57.1%; Pred. No. 7e+02;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
 |||:
 Db 96 CSPLSGC 102

RESULT 35
 ID AAO00015
 AA000015 standard; protein; 193 AA.

XX AC AA000015;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 13907.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KM tissue growth factor; immunomodulatory; cancer; leukaemia;
 KM nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004927.

XX PR 28-FEB-2000; 2000US-00515126.

XX PR 18-MAY-2000; 2000US-00577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-514838/56.

XX DR N-PSDB; AA179946.

XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 XX and treating e.g. leukemia, inflammation and immune disorders.

XX PS Claim 20; SEQ ID NO 13907; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 193 AA;

Query Match 70.5%; Score 31; DB 4; Length 193;
 Best Local Similarity 57.1%; Pred. No. 7.3e+02;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
 |||:
 Db 98 CSPLSGC 104

RESULT 36
 ID ADF60298
 ADF60298 standard; protein; 193 AA.

XX AC ADF60298;

XX DT 12-FEB-2004 (first entry)

XX DE Human contig polypeptide sequence SEQ ID NO:2665.

XX biological activity; genetic engineering; hybridisation probe; oligomer;
 KM primer; chromosome mapping; gene mapping; recombinant protein production;
 KM human.

XX OS Homo sapiens.

XX PN WO2003080795-A2.

XX PD 02-OCT-2003.

XX PF 09-AUG-2002; 2002WO-US025485.

XX PR 09-AUG-2001; 2001US-0311261P.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Yang Y, Wang Z, Weng G, Ma Y;

XX DR WPI: 2003-876918/81.

XX DR N-PSDB; ADF59846.

XX PT New polynucleotides, useful as hybridization probes, oligomers or
 XX primers, for chromosome or gene mapping, for the recombinant production
 XX of proteins, and for generating antisense DNA or RNA.

XX PS Example 3; SEQ ID NO 2665; 571pp; English.

XX CC The present invention describes isolated polynucleotide sequences (I),
 CC which encode polypeptides (II) with biological activity. Also described:
 CC (1) a vector comprising (I); (2) an expression vector comprising (I); (3)
 CC a host cell genetically engineered to comprise (I) which is operatively
 CC associated with a regulatory sequence that modulates expression of (I) in
 CC the host cell; (4) a polypeptide (II) encoded by (I); (5) a composition
 CC comprising the polypeptide of (4) and a carrier; (6) an antibody directed
 CC against the polypeptide of (4); (7) detecting (I) or the polypeptide of
 CC (4) in a sample; (8) identifying a compound that binds to the polypeptide
 CC of (4); (9) producing the polypeptide of (4); and (10) a collection of
 CC polynucleotides comprising at least one of the polynucleotide sequences
 CC (I). The polynucleotides (I) can be used as hybridisation probes,
 CC oligomers or primers, for chromosome or gene mapping, for the recombinant
 CC production of proteins, and for generating antisense DNA or RNA. The
 CC present sequence represents a human contig polypeptide sequence, which is
 CC used in an example from the present invention.

XX SQ Sequence 193 AA;

Query Match 70.5%; Score 31; DB 7; Length 193;
 Best Local Similarity 57.1%; Pred. No. 7.3e+02;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
 |||:
 Db 98 CSPLSGC 104

RESULT 37
 ID ABP9259
 ABP9259 standard; protein; 238 AA.

```

XX AC ABP99259;
XX XX
XX DT 21-MAR-2003 (first entry)
XX DE Orthosomycin biosynthetic polypeptide SEQ ID NO 105.
XX KW Orthosomycin; biosynthesis; everninomicin; avilamycin; enzyme.
XX OS Streptomyces mobaraensis.
XX PN WO200279505-A2.
XX PD 10-OCT-2002.
XX PF 28-MAR-2002; 2002WO-CA000432.
XX PR 28-MAR-2001; 2001US-0279095P.
XX PR 30-MAR-2001; 2001US-0279709P.
XX PR 20-APR-2001; 2001US-0285214P.
XX PS (ECOP-) ECOPIA BIOSCIENCES INC.
XX PI Farnet CM, Zazopoulos E, Staffa A;
XX DR WPI; 2003-058435/05.
XX DR N-PSDB; ABZ66722.
XX PT Identifying orthosomycin biosynthetic gene, gene fragment or gene
XX PT cluster, by detecting presence of nucleic acid sequence corresponding to
XX PT 17 of flamamycins protein families.
XX PS Claim 1; Page 260-261, 511pp; English.
XX CC The invention relates to identifying orthosomycin biosynthetic genes and
XX CC its fragment/gene cluster (ABZ6670-ABZ66813), comprising detecting the
XX CC presence of a nucleic acid sequence coding for a polypeptide (ABP99207-
XX CC ABP99362). The method is useful for identifying an orthosomycin
XX CC biosynthetic gene, gene fragment or gene cluster, especially an
XX CC everninomicin-type or avilamycin-type orthosomycin biosynthetic gene,
XX CC gene fragment or gene cluster. The method is useful for detecting the
XX CC presence of any organism that contains DNA for the production of
XX CC orthosomycins (both everninomicin-type orthosomycins and avilamycin-type
XX CC orthosomycins) regardless of the level at which genes for orthosomycin
XX CC production are expressed by the organism or the amount of orthosomycin
XX CC produced by the organism. This allows for the detection of new
XX CC orthosomycin natural products, not produced by the organism
XX CC
XX SQ Sequence 238 AA;
XX
XX Query Match 70.5%; Score 31; DB 6; Length 238;
XX Best Local Similarity 57.1%; Pred. No. 8.7e+02;
XX Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 4 CXPXTGC 10
XX | | : | |
XX DB 26 CAPDSCG 32
XX
XX RESULT 38
XX ABP99348
XX ID ABP99348 standard; protein; 239 AA.
XX XX
XX AC ABP99348;
XX XX
XX DT 21-MAR-2003 (first entry)
XX DE Orthosomycin biosynthetic polypeptide GenBank No AAK83184.
XX KW Orthosomycin; biosynthesis; everninomicin; avilamycin; enzyme.
XX OS Streptomyces viridochromogenes.
XX XX

```

```

PN WO200279505-A2.
XX XX
XX PD 10-OCT-2002.
XX XX
XX PF 28-MAR-2002; 2002WO-CA000432.
XX PR 28-MAR-2001; 2001US-0279095P.
XX PR 30-MAR-2001; 2001US-0279709P.
XX PR 20-APR-2001; 2001US-0285214P.
XX PS (ECOP-) ECOPIA BIOSCIENCES INC.
XX PI Farnet CM, Zazopoulos E, Staffa A;
XX DR WPI; 2003-058435/05.
XX PT Identifying orthosomycin biosynthetic gene, gene fragment or gene
XX PT cluster, by detecting presence of nucleic acid sequence corresponding to
XX PT 17 of flamamycins protein families.
XX PS Claim 1; Page; 511pp; English.
XX CC The invention relates to identifying orthosomycin biosynthetic genes and
XX CC its fragment/gene cluster (ABZ6670-ABZ66813), comprising detecting the
XX CC presence of a nucleic acid sequence coding for a polypeptide (ABP99207-
XX CC ABP99362). The method is useful for identifying an orthosomycin
XX CC biosynthetic gene, gene fragment or gene cluster, especially an
XX CC everninomicin-type or avilamycin-type orthosomycin biosynthetic gene,
XX CC gene fragment or gene cluster. The method is useful for detecting the
XX CC presence of any organism that contains DNA for the production of
XX CC orthosomycins (both everninomicin-type orthosomycins and avilamycin-type
XX CC orthosomycins) regardless of the level at which genes for orthosomycin
XX CC production are expressed by the organism or the amount of orthosomycin
XX CC produced by the organism. This allows for the detection of new
XX CC orthosomycin natural products, not produced by the organism. Note: The
XX CC present sequence is not given in the specification but was obtained from
XX CC GenBank
XX SQ Sequence 239 AA;
XX
XX Query Match 70.5%; Score 31; DB 6; Length 239;
XX Best Local Similarity 57.1%; Pred. No. 8.7e+02;
XX Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 4 CXPXTGC 10
XX | | : | |
XX DB 22 CAPDSCG 28
XX
XX RESULT 39
XX ABP76713
XX ID ABP76713 standard; protein; 239 AA.
XX XX
XX AC ABP76713;
XX XX
XX DT 26-FEB-2003 (first entry)
XX DE Streptomyces viridochromogenes AV166.
XX KW Avilamycin; antibacterial; virucide; protozoacide; fungicide; infection;
XX KW medicine; Staphylococcus aureus; biosynthetic gene cluster; enzyme.
XX OS Streptomyces viridochromogenes.
XX PN WO200268436-A1.
XX PD 06-SEP-2002.
XX PF 24-AUG-2001; 2001WO-BP009815.
XX PR 25-FEB-2001; 2001DE-01009166.
XX PS (COMB-) COMBINATURE BIOPHARM AG.
XX XX

```

```
XX
PI Weinauer G, Muehlenweg A, Treifzer A, Bechtold A;
XX
DR WPI; 2003-018650/01.
DR N-PSDB; AB237547.
XX
PT New avilamycin derivatives, useful for treatment of infections, and
PT nucleic acid encoding avilamycin synthesis enzymes.
XX
PS Claim 16; Page 187-189; 31ppp; German.
XX
CC The invention relates to avilamycin derivatives (I) with antibacterial,
CC virucide, protozoacide and fungicide activity. (I) are useful for
CC treatment of infections (bacterial, viral, protozoal or fungal), in human
CC or veterinary medicine, particularly where caused by Staphylococcus
CC aureus. (I) are more hydrophilic than known avilamycins. The present
CC sequence is that of an avilamycin synthesis enzyme from the Streptomyces
CC viridochromogenes Avilamycin A biosynthetic gene cluster (AB237515-
CC AB237516)
CC
SQ Sequence 239 AA;
XX
Query Match 70.5%; Score 31; DB 6; Length 239;
Best Local Similarity 57.1%; Pred. No. 8.7e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
QY 4 CXPXTGC 10
DB 22 CAPDSCG 28
XX
RESULT 40
ADX95230
ID ADX95230 standard; protein; 285 AA.
XX
AC ADX95230;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polypeptide seqid 57894.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
XX
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIU/J) LIU J.
PA (ZHOU/J) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABAS/) TABASKA J E.
PA (CAOV/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
DR WPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
```

```
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 57894; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 285 AA;
XX
Query Match 70.5%; Score 31; DB 8; Length 285;
Best Local Similarity 57.1%; Pred. No. 1e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
QY 4 CXPXTGC 10
DB 160 CAPLSCG 166
XX
RESULT 41
ABO74638
ID ABO74638 standard; protein; 310 AA.
XX
AC ABO74638;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #6813.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US651795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, DeLoughery C, Bush D;
XX
DR WPI; 2003-615309/58.
XX
DR N-PSDB; ABD08209.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure, SEQ ID NO 23384; 45pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
```


CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-*P. aeruginosa*-derived peptides or polypeptides, as target
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology. Sequences ABO67826-
CC ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
CC
SQ Sequence 310 AA;

Query Match 70.5%; Score 31; DB 7; Length 310;
Best Local Similarity 57.1%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 236 CSPAAGC 242

RESULT 42
ADU17203
ID ADU17203 standard; protein; 354 AA.
XX
XX
AC ADU17203;
XX
DT 13-JAN-2005 (first entry)
XX
DE M. tuberculosis exported polypeptides #611.
XX
XX M. tuberculosis exported polypeptides #611.
XX
XX anti-tubercular; tuberculostatic; vaccine; tuberculosis; hepatitis.
XX
OS Mycobacterium tuberculosis.
XX
XX
XX US2004214165-A1.
XX
XX
XX 28-OCT-2004.
XX
XX 16-MAY-2001; 2001US-00855604.
XX
XX 16-MAY-2001; 2001US-00855604.
XX
XX
XX (GIGO/) GICQUEL B.
XX (PORT/) PORTNOI D.
XX (LIME/) LIM E.
XX (PELI/) PELICIC V.
XX (GUIG/) GUIGUENO A.
XX (GOGU/) GOGUET D L S Y.
XX
XX
XX Gicquel B, Portnoi D, Lim E, Pelicic V, Guigueno A, Goguett DLSY;
XX
XX WPI; 2004-765598/75.
XX N-PSDB; ADU17188.
XX
XX
XX Novel recombinant screening, cloning or expression vector replicating in
XX mycobacteria, useful for preventing Mycobacterium tuberculosis, having
XX replicon that is functional in mycobacteria, selectable marker and
XX reporter cassette.
XX
XX Example 5; SEQ ID NO 863; 220bp; English.
XX
XX
XX The invention relates to a recombinant screening, cloning or expression
XX vector, that replicates in mycobacteria, comprising a replicon that is
XX functional in mycobacteria, selectable marker, reporter cassette. The
XX recombinant vector is useful for diagnosing and preventing M.
XX tuberculosis. Also disclosed is a vaccine composition for tuberculosis

CC and hepatitis. The present sequence represents the amino acid sequence of
CC a Mycobacterium tuberculosis exported polypeptide.
XX
SQ Sequence 354 AA;

Query Match 70.5%; Score 31; DB 8; Length 354;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 80 CSPAVGC 86

RESULT 43
ABO70819
ID ABO70819 standard; protein; 444 AA.
XX
XX
AC ABO70819;
XX
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #2994.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
XX
XX US6551795-B1.
XX
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
XX WPI; 2003-615309/58.
XX N-PSDB; ABO4390.
XX
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 19565; 455bp; English.
XX
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a *P. aeruginosa* nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-*P. aeruginosa* drugs, as templates for recombinant
XX production of *P. aeruginosa*-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of *P. aeruginosa*-caused
XX infection, and in detection of *P. aeruginosa* sequences or other sequences
XX of *Pseudomonas* species using biochip technology. Sequences ABO67826-
XX ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX
XX Sequence 444 AA;

Query Match 70.5%; Score 31; DB 7; Length 444;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
 DB 47 CSPAAAGC 53

RESULT 44

ABB93340
 ID ABB93340 standard; protein; 482 AA.

AC ABB93340;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 2551.

KM Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP009892.

PR 28-AUG-2001; 2001WO-EP009892.

PA (FARB) BAYER AG.

PI Tietjen K, Weidler M;

DR WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,
 comprising aligning and comparing nucleic acid or amino acid sequences
 from plant with nucleic acid or amino acid sequences from non-plant
 organisms.

PS Claim 5; SEQ ID NO 2551; 261pp + Sequence Listing; English.

CC The invention relates to identifying target proteins (ABB90790-ABB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for
 CC identifying modulators. The identified modulators are useful as
 CC herbicides

SQ Sequence 482 AA;

Query Match 70.5%; Score 31; DB 5; Length 482;

Best Local Similarity 57.1%; Pred. No. 1.6e+03; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10

DB 301 CTPGSGC 307

RESULT 45

AAW32363
 ID AAW32363 standard; protein; 580 AA.

AC AAW32363;

DT 13-JAN-1998 (first entry)

DE Mycobacterium tuberculosis antigen TBRa29.

KM Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;

KW skin testing; M.tuberculosis.

XX Mycobacterium tuberculosis.

OS WO9709429-A2.

PN 13-MAR-1997.

PF 30-AUG-1996; 96WO-US014675.

PR 01-SEP-1995; 95US-00523435.

PR 22-SEP-1995; 95US-00532136.

PR 22-MAR-1996; 96US-00620280.

PR 05-JUN-1996; 96US-00658800.

PR 12-JUL-1996; 96US-00680573.

PA (CORI-) CORIXA CORP.

PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;

PI Vedvick TH, Twardzik DR;

DR WPI; 1997-192904/17.

DR N-PSDB; AAT91410.

PT New immunogenic polypeptide(s) from soluble M. tuberculosis antigens -

PT useful for diagnosis of M. tuberculosis infection.

PS Example 3; Page 118-121, 190pp; English.

CC A new immunogenic polypeptide has been developed comprising an
 CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its
 CC variant differing only in conservative substitutions and/or
 CC modifications). The present sequence represents a M.tuberculosis antigen,
 CC TBRa29. The immunogenic polypeptide can be used to diagnose
 CC M.tuberculosis infection by forming complexes with specific antibodies in
 CC the sample. Fragments of DNA encoding the immunogenic polypeptide can be
 CC used as diagnostic primers or probes and agents that bind to the antigen,
 CC especially monoclonal antibodies or equivalent polyclonal antibodies, are
 CC also used for diagnosis

SQ Sequence 580 AA;

Query Match 70.5%; Score 31; DB 2; Length 580;

Best Local Similarity 57.1%; Pred. No. 1.8e+03; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10

DB 58 CSPAVGC 64

RESULT 46

AAW32431
 ID AAW32431 standard; protein; 580 AA.

AC AAW32431;

DT 08-JAN-1998 (first entry)

DE Mycobacterium tuberculosis antigen TBRa29.

KM Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;

KW skin testing; M.tuberculosis.

OS Mycobacterium tuberculosis.

PN WO9709428-A2.

PD 13-MAR-1997.

PF 30-AUG-1996; 96WO-US014674.

PR 01-SEP-1995; 95US-00523436.

PR 22-SEP-1995; 95US-00533634.

22-MAR-1996; 96US-00620874.
 PR 05-JUN-1996; 96US-00659683.
 PR 12-JUL-1996; 96US-00680574.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Reed SG, Skeiky YA, Dillon DC, Campos-Neto A, Houghton R;
 PI Vedvick TS, Twardzik DR;
 XX
 DR MPI: 1997-192903/17.
 DR N-PSDB; AAT91473.
 XX
 PT New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are
 PT useful in vaccines for prevention or treatment of tuberculosis, also for
 PT diagnosis.
 XX
 PS Example 3; Page 109-111; 168pp; English.
 XX
 CC A new immunogenic polypeptide has been developed comprising an
 CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its
 CC variant differing only in conservative substitutions and/or
 CC modifications). The present sequence represents a M.tuberculosis antigen,
 CC TbrA29. The immunogenic protein, and fusion proteins containing one or
 CC more of the proteins or one of the proteins plus ESAT-6, are useful in
 CC vaccines, preferably when formulated with a non-specific adjuvant, to
 CC induce an immune response against M.tuberculosis (for treatment or
 CC prevention)
 CC
 SQ Sequence 580 AA;

Query Match 70.5%; Score 31; DB 2; Length 580;
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
 | | | |
 DB 58 CSPAVGC 64

RESULT 47
 ID AAM64303
 AC AAM64303;
 XX
 DT 17-OCT-2003 (revised)
 DT 09-NOV-1998 (first entry)
 XX
 DB Mycobacterium tuberculosis antigen TbrA29.
 XX
 KW Tuberculosis; infection; diagnosis; antigen; TbrA29.
 XX
 OS Mycobacterium tuberculosis; strain H37Ra.
 XX
 PN WO9816645-A2.
 XX
 PD 23-APR-1998.
 XX
 PF 07-OCT-1997; 97WO-US018214.
 XX
 PR 11-OCT-1996; 96US-00729622.
 PR 13-MAR-1997; 97US-00818111.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
 PI Vedvick TS, Twardzik DR, Lodes MJ;
 XX
 DR MPI: 1998-251292/22.
 DR N-PSDB; AAV44351.
 XX
 PT New isolated Mycobacterium tuberculosis polypeptides and DNA - used to
 PT develop products for the detection of M. tuberculosis infection and

PT diagnosis of tuberculosis.
 XX
 PS Example 3; Page 111-113; 250pp; English.
 XX

CC This polypeptide comprises Mycobacterium tuberculosis soluble antigen
 CC TbrA29. It is encoded by a DNA sequence (see AAV44351) isolated from a M.
 CC tuberculosis strain H37Ra expression library with rabbit anti-sera raised
 CC against M. tuberculosis supernatant. TbrA29 shows some homology to a
 CC sequence previously identified in Mycobacterium leprae sequence but not
 CC in M. tuberculosis. The invention relates to compositions and methods for
 CC diagnosing tuberculosis. It provides polypeptides (see AAM64291-W64379)
 CC comprising an antigenic portion of a soluble M. tuberculosis antigen, or
 CC an immunogenic portion of an M. tuberculosis antigen, as well as DNA
 CC sequences encoding such polypeptides, recombinant expression vectors and
 CC transformed or transfected host cells. Also claimed are methods and
 CC diagnostic kits for detecting M. tuberculosis infection in a patient
 CC using the above polypeptides, antibodies or oligonucleotide probes and
 CC primers, for the diagnosis of tuberculosis. (Updated on 17-OCT-2003 to
 CC standardise OS field)

Query Match 70.5%; Score 31; DB 2; Length 580;
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
 | | | |
 DB 58 CSPAVGC 64

RESULT 48
 ID AAM81666
 AC AAM81666;
 XX
 DT 27-JAN-1999 (first entry)
 XX
 DE M. tuberculosis immunogenic polypeptide TbrA29.
 XX
 KW Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
 KW vaccine; pharmaceutical; infection; diagnosis.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN WO9816646-A2.
 XX
 PD 23-APR-1998.
 XX
 PF 07-OCT-1997; 97WO-US018293.
 XX
 PR 11-OCT-1996; 96US-00730510.
 PR 13-MAR-1997; 97US-00818112.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
 PI Vedvick TS, Twardzik DR, Lodes MJ;
 XX
 DR MPI: 1998-261042/23.
 DR N-PSDB; AAV64459.
 XX
 PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used to
 PT develop products for the detection of M. tuberculosis infection and for
 PT diagnosis, treatment and prevention of tuberculosis.
 XX
 PS Example 3; Page 107-108; 230pp; English.
 XX
 CC This sequence represents an immunogenic portion of a soluble
 CC Mycobacterium tuberculosis (MT) antigen which can be used in a method for
 CC inducing protective immunity against tuberculosis (TB). This sequence can
 CC be formulated into vaccines and/or pharmaceutical compositions for

CC immunising against M. tuberculosis infection or may be used for the
 CC diagnosis of tuberculosis
 XX
 SQ Sequence 580 AA;

Query Match 70.5%; Score 31; DB 2; Length 580;
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
 | | | |
 DB 58 CSPAVGC 64

RESULT 49
 AAY38968
 ID AAY38968 standard; protein; 580 AA.

XX AAY38968;

XX 05-NOV-1999 (first entry)

XX M. tuberculosis recombinant antigen protein TBra29.

XX Antigen; diagnosis; detection; infection; antibody; immunisation;
 KW vaccine; immunity.

OS Mycobacterium tuberculosis.

PN WO9942118-A2.

XX 26-AUG-1999.

XX 17-FEB-1999; 99WO-US003265.

XX 18-FEB-1998; 98US-00024753.

XX 05-MAY-1998; 98US-00072596.

XX (CORI-) CORIXA CORP.

XX Reed SG, Skeiky YAM, Dillon DC, Campos-Neto A, Houghton R;
 PI Veddyck TS, Twardzik DR, Lodes MJ, Hendrickson RC;

XX WPI: 1999-527416/44.

XX N-PSDB; AA219049.

XX New polypeptide comprising antigenic portions of M. tuberculosis.

XX Example 3; Page 147-149; 323pp; English.

CC This invention describes novel recombinant antigens and their encoding
 CC nucleic acids derived from Mycobacterium tuberculosis. The novel
 CC polypeptides are useful for detecting M. tuberculosis infection in a
 CC biological sample by detecting antibodies which bind with the
 CC polypeptides, and are useful as vaccines for immunizing against M.
 CC tuberculosis infection. The new detection methods are needed as current
 CC vaccination strategies do not provide 100% immunity

XX Sequence 580 AA;

Query Match 70.5%; Score 31; DB 2; Length 580;
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
 | | | |
 DB 58 CSPAVGC 64

RESULT 50
 AAY39105
 ID AAY39105 standard; protein; 580 AA.

AC AAY39105;
 XX
 DT 05-NOV-1999 (first entry)

XX M. tuberculosis antigen TBra29 amino acid sequence.

XX Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
 KW immunotherapy; diagnosis; immunisation; vaccine; infection;
 KW immune response; skin test.

OS Mycobacterium tuberculosis.

PN WO9942076-A2.

XX 26-AUG-1999.

XX 17-FEB-1999; 99WO-US003268.

XX 18-FEB-1998; 98US-00025197.

XX 05-MAY-1998; 98US-00072967.

XX (CORI-) CORIXA CORP.

XX Reed SG, Skeiky YAM, Dillon DC, Campos-Neto A, Houghton R;
 PI Veddyck TS, Twardzik DR, Lodes MJ, Hendrickson RC;

XX WPI: 1999-527409/44.

XX N-PSDB; AA219261.

XX New antigens from Mycobacterium tuberculosis useful in diagnostic skin
 PT tests and protective or therapeutic vaccines or compositions.
 XX
 PS Example 3; Page 105-106; 299pp; English.

CC The present invention describes polypeptides comprising an immunogenic
 CC part of a Mycobacterium tuberculosis antigen (Ag). Also described are
 CC vaccines and fusion protein containing M. tuberculosis Ag's. M.
 CC tuberculosis Ag's, DNAs encoding them, derived fusion proteins and other
 CC polypeptides fragments, can be used in pharmaceutical compositions or
 CC vaccines to generate a protective or therapeutic immune response to M.
 CC tuberculosis and as reagents in skin tests for diagnosis of tuberculosis.
 CC Ag can induce proliferation of, or cytokine secretion by, T, B or natural
 CC killer cells and/or macrophages in tuberculosis-immune subjects. AA219249
 CC to AA219460 and AAY39083 to AAY39225 are used in the exemplification of
 CC the present invention

XX Sequence 580 AA;

Query Match 70.5%; Score 31; DB 2; Length 580;
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
 | | | |
 DB 58 CSPAVGC 64

Search completed: January 4, 2006, 15:55:15
 Job time : 128.965 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 ; Search time 95.8087 Seconds

(without alignments)
64.204 Million cell updates/sec

Title: US-09-932-322-2

Perfect score: 30

Sequence: 1 XXXXXXXXXXXXXXX 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database :

A_Geneseq_21:.*
1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*
9: geneseqp2005s:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	60.0	9	AD013022	Adq13022 Hepatitis
2	18	60.0	8	AD013021	Adq13021 Hepatitis
3	18	60.0	11	AAW5943	AAW5943 Molecule
4	18	60.0	11	AAW5944	AAW5944 Molecule
5	18	60.0	15	ADV23688	Adv23688 HBV Immun
6	18	60.0	15	ADV23687	Adv23687 HBV Immun
7	18	60.0	24	ABBA2273	Abba2273 Peptide #
8	18	60.0	24	AAW36081	AAW36081 Peptide #
9	18	60.0	24	AAW75973	AAW75973 Human bon
10	18	60.0	24	AAW63159	AAW63159 Human bon
11	18	60.0	24	ABG57702	Abg57702 Human liv
12	18	60.0	69	ABBB6895	Abbb6895 Drosophil
13	18	60.0	69	ADP30715	Adp30715 Human sec
14	18	60.0	91	ABR39058	ABR39058 Peptide #
15	18	60.0	91	AAW32549	AAW32549 Peptide #
16	18	60.0	91	AAW72290	AAW72290 Human bon
17	18	60.0	91	AAW59710	AAW59710 Human bon
18	18	60.0	91	ABG53976	Abg53976 Human liv
19	18	60.0	91	ABG42105	Abg42105 Human liv
20	18	60.0	93	ADP30859	Adp30859 Human sec
21	18	60.0	99	ABO83167	ABO83167 Pseudomon
22	18	60.0	99	ADP30709	Adp30709 Human sec
23	18	60.0	107	ABO74878	ABO74878 Pseudomon
24	18	60.0	115	ABO73394	ABO73394 Pseudomon

25	18	60.0	122	7	ADP04777	Adp04777 Bacterial
26	18	60.0	135	7	ADP30703	Adp30703 Human sec
27	18	60.0	144	7	ABO75730	ABO75730 Pseudomon
28	18	60.0	144	8	ADP31453	Adp31453 Human sec
29	18	60.0	150	8	ADP30450	Adp30450 Plant ful
30	18	60.0	150	6	ABU45351	Abu45351 Protein e
31	18	60.0	153	3	AAAB1043	AAAB1043 Human ORF
32	18	60.0	153	5	ABP09251	ABP09251 Human ORF
33	18	60.0	154	7	ABO74421	ABO74421 Pseudomon
34	18	60.0	156	7	ABO77836	ABO77836 Pseudomon
35	18	60.0	159	7	ADP30761	Adp30761 Human sec
36	18	60.0	162	8	ADP30758	Adp30758 Human sec
37	18	60.0	162	8	ADP30756	Adp30756 Human sec
38	18	60.0	162	8	ADP30757	Adp30757 Human sec
39	18	60.0	162	8	ADP30813	Adp30813 Human sec
40	18	60.0	165	8	ADP31269	Adp31269 Human sec
41	18	60.0	168	8	ADP31099	Adp31099 Human sec
42	18	60.0	171	8	ADP30794	Adp30794 Human sec
43	18	60.0	171	8	ADP30793	Adp30793 Human sec
44	18	60.0	177	3	AAG10193	AAG10193 Arabidops
45	18	60.0	177	8	ADP30755	Adp30755 Human sec
46	18	60.0	177	8	ADP31682	Adp31682 Human sec
47	18	60.0	177	8	ADP36379	Adp36379 Plant pol
48	18	60.0	182	7	ABM86593	ABM86593 Rice abio
49	18	60.0	185	7	ABM89559	ABM89559 Rice abio
50	18	60.0	192	8	ADP31576	Adp31576 Human sec
51	18	60.0	194	7	ABO78166	ABO78166 Pseudomon
52	18	60.0	199	8	ADY13615	ADY13615 Plant ful
53	18	60.0	201	8	ADP30711	Adp30711 Human sec
54	18	60.0	204	3	AAG10192	AAG10192 Arabidops
55	18	60.0	210	8	ADP30986	Adp30986 Human sec
56	18	60.0	210	8	ADP30960	Adp30960 Human sec
57	18	60.0	219	8	ADP30593	Adp30593 Human sec
58	18	60.0	222	8	ADP31354	Adp31354 Human sec
59	18	60.0	228	8	ADP30764	Adp30764 Human sec
60	18	60.0	233	3	AAV74791	AAV74791 Neisseria
61	18	60.0	234	8	ADP31480	Adp31480 Human sec
62	18	60.0	234	8	ADP31481	Adp31481 Human sec
63	18	60.0	237	8	ADP30792	Adp30792 Human sec
64	18	60.0	239	7	ABO81291	ABO81291 Pseudomon
65	18	60.0	249	8	ADP30754	Adp30754 Human sec
66	18	60.0	249	8	ADP31443	Adp31443 Human sec
67	18	60.0	252	8	ADP31485	Adp31485 Human sec
68	18	60.0	258	7	ABO80689	ABO80689 Pseudomon
69	18	60.0	258	8	ADP30597	Adp30597 Human sec
70	18	60.0	267	8	ADP30822	Adp30822 Human sec
71	18	60.0	270	7	ABO68843	ABO68843 Pseudomon
72	18	60.0	276	8	ADP30568	Adp30568 Human sec
73	18	60.0	277	8	ADP30503	Adp30503 Human sec
74	18	60.0	279	8	ADP31107	Adp31107 Human sec
75	18	60.0	291	8	ADP31193	Adp31193 Human sec
76	18	60.0	294	8	ADP31045	Adp31045 Human sec
77	18	60.0	300	8	ADP30775	Adp30775 Human sec
78	18	60.0	309	8	ADP30862	Adp30862 Human sec
79	18	60.0	309	8	ADP30873	Adp30873 Human sec
80	18	60.0	310	7	ABO80230	ABO80230 Pseudomon
81	18	60.0	312	8	ADP31454	Adp31454 Human sec
82	18	60.0	318	7	ABM85364	ABM85364 Mouse pro
83	18	60.0	318	8	ADP30588	Adp30588 Human sec
84	18	60.0	320	8	ADP30824	Adp30824 Human sec
85	18	60.0	320	8	ADP31607	Adp31607 Human sec
86	18	60.0	320	8	ADP31649	Adp31649 Human sec
87	18	60.0	323	7	ABO74543	ABO74543 Pseudomon
88	18	60.0	327	7	ABM89959	ABM89959 Rice abio
89	18	60.0	330	4	ABBS9598	ABBS9598 Drosophil
90	18	60.0	339	8	ADP30892	Adp30892 Human sec
91	18	60.0	339	8	ADP30702	Adp30702 Human sec
92	18	60.0	345	8	ADP31016	Adp31016 Human sec
93	18	60.0	345	8	ADP31683	Adp31683 Human sec
94	18	60.0	346	8	ADP38850	ADP38850 Plant pol
95	18	60.0	348	8	ADP31676	Adp31676 Human sec
96	18	60.0	348	8	ADP31374	Adp31374 Human sec
97	18	60.0	353	8	ADP31558	Adp31558 Human sec

98	18	60.0	354	8	ADP30783	Adp30783	Human	sec	171	18	60.0	558	8	ADP31255	Adp31255	Human	sec
99	18	60.0	354	8	ADP30707	Adp30707	Human	sec	172	18	60.0	564	8	ADP31194	Adp31194	Human	sec
100	18	60.0	354	8	ADP30779	Adp30779	Human	sec	173	18	60.0	576	8	ADP31133	Adp31133	Human	sec
101	18	60.0	354	8	ADP30784	Adp30784	Human	sec	174	18	60.0	574	8	ADP31248	Adp31248	Human	sec
102	18	60.0	354	8	ADP30778	Adp30778	Human	sec	175	18	60.0	583	2	ADH11358	Vertebrat		
103	18	60.0	355	4	ABM57739	Abm57739	Drosophil		176	18	60.0	588	8	ADP31680	Adp31680	Human	sec
104	18	60.0	355	7	ABM87551	Abm87551	Rice	abio	177	18	60.0	592	8	ADP30917	Adp30917	Human	sec
105	18	60.0	355	9	ADZ76101	Adz76101	Fruit	fly	178	18	60.0	598	8	ADP30895	Adp30895	Human	sec
106	18	60.0	357	9	ADP31223	Adp31223	Human	sec	179	18	60.0	603	8	ADP31150	Adp31150	Human	sec
107	18	60.0	357	8	ADP31267	Adp31267	Human	sec	180	18	60.0	604	8	ADP30940	Adp30940	Human	sec
108	18	60.0	360	8	ADP31439	Adp31439	Human	sec	181	18	60.0	605	8	ADP30941	Adp30941	Human	sec
109	18	60.0	366	8	ADP31106	Adp31106	Human	sec	182	18	60.0	604	8	ADP30507	Adp30507	Human	sec
110	18	60.0	366	8	ADP31670	Adp31670	Human	sec	183	18	60.0	605	8	ADP31263	Adp31263	Human	sec
111	18	60.0	367	5	AAO21694	Aao21694	Human	sec	184	18	60.0	609	8	ADP31363	Adp31363	Human	sec
112	18	60.0	369	7	ABM8560	Abm8560	Rice	abio	185	18	60.0	612	8	ADP31064	Adp31064	Human	sec
113	18	60.0	370	8	ADY09767	Ady09767	Plant	full	186	18	60.0	615	8	ADP31132	Adp31132	Human	sec
114	18	60.0	372	8	ADP31610	Adp31610	Human	sec	187	18	60.0	615	8	ADP31361	Adp31361	Human	sec
115	18	60.0	382	8	ADP31216	Adp31216	Human	sec	188	18	60.0	617	8	ADP31360	Adp31360	Human	sec
116	18	60.0	382	8	ADP31221	Adp31221	Human	sec	189	18	60.0	617	8	ADP31657	Adp31657	Human	sec
117	18	60.0	387	8	ADP31158	Adp31158	Human	sec	190	18	60.0	618	8	ADP31554	Adp31554	Human	sec
118	18	60.0	390	8	ADP31158	Adp31158	Human	sec	191	18	60.0	621	8	ADP30896	Adp30896	Human	sec
119	18	60.0	397	8	ADP31218	Adp31218	Human	sec	192	18	60.0	639	8	ADP31521	Adp31521	Human	sec
120	18	60.0	399	8	ADP31345	Adp31345	Human	sec	193	18	60.0	639	8	ADP30861	Adp30861	Human	sec
121	18	60.0	401	7	ABO68344	AbO68344	Pseudomon		194	18	60.0	642	8	ADP31265	Adp31265	Human	sec
122	18	60.0	411	8	ADP31104	Adp31104	Human	sec	195	18	60.0	645	8	ADP31124	Adp31124	Human	sec
123	18	60.0	411	8	ADP30729	Adp30729	Human	sec	196	18	60.0	645	8	ADP31125	Adp31125	Human	sec
124	18	60.0	420	8	ADP31349	Adp31349	Human	sec	197	18	60.0	645	8	ADP30858	Adp30858	Human	sec
125	18	60.0	421	8	ADP31159	Adp31159	Human	sec	198	18	60.0	645	8	ADP31183	Adp31183	Human	sec
126	18	60.0	423	8	ADP31479	Adp31479	Human	sec	199	18	60.0	654	8	ADP30666	Adp30666	Human	sec
127	18	60.0	425	8	ADP30549	Adp30549	Human	sec	200	18	60.0	658	8	ADP31226	Adp31226	Human	sec
128	18	60.0	429	8	ADP31207	Adp31207	Human	sec	201	18	60.0	665	8	ADP30571	Adp30571	Human	sec
129	18	60.0	434	5	ABP41972	Abp41972	Human	Ova	202	18	60.0	666	8	ADP31547	Adp31547	Human	sec
130	18	60.0	441	8	ADP31206	Adp31206	Human	sec	203	18	60.0	666	8	ADP31256	Adp31256	Human	sec
131	18	60.0	449	8	ADQ21283	Adq21283	Human	sof	204	18	60.0	669	8	ADP31598	Adp31598	Human	sec
132	18	60.0	450	8	ADP30887	Adp30887	Human	sec	205	18	60.0	669	8	ADP31493	Adp31493	Human	sec
133	18	60.0	453	8	ADP31465	Adp31465	Human	sec	206	18	60.0	677	8	ADP31585	Adp31585	Human	sec
134	18	60.0	471	8	ADP31567	Adp31567	Human	sec	207	18	60.0	681	8	ADP30868	Adp30868	Human	sec
135	18	60.0	478	8	ADP31007	Adp31007	Human	sec	208	18	60.0	681	8	ADP31053	Adp31053	Human	sec
136	18	60.0	483	8	ADP30852	Adp30852	Human	sec	209	18	60.0	699	8	ADP31080	Adp31080	Human	sec
137	18	60.0	483	8	ADP30853	Adp30853	Human	sec	210	18	60.0	699	8	ADP31079	Adp31079	Human	sec
138	18	60.0	492	4	AA871707	Aa871707	Human	ERR	211	18	60.0	699	8	ADP31241	Adp31241	Human	sec
139	18	60.0	492	7	ADDI4170	AdDi4170	Human	src	212	18	60.0	699	8	ADP31240	Adp31240	Human	sec
140	18	60.0	494	6	ABR64215	AbR64215	Angiogene		213	18	60.0	702	8	ADP31518	Adp31518	Human	sec
141	18	60.0	494	8	ADO19178	Ado19178	Human	PRO	214	18	60.0	711	8	ADP31215	Adp31215	Human	sec
142	18	60.0	494	8	ADO19174	Ado19174	Human	PRO	215	18	60.0	711	8	ADP31552	Adp31552	Human	sec
143	18	60.0	494	8	ADP54221	Adp54221	Human	PRO	216	18	60.0	711	8	ADP31535	Adp31535	Human	sec
144	18	60.0	494	9	ADPX05928	Adpx05928	Cyclin-de		217	18	60.0	711	8	ADP31440	Adp31440	Human	sec
145	18	60.0	494	9	ADY14502	Ady14502	PRO	polyp	218	18	60.0	714	8	ADP31561	Adp31561	Human	sec
146	18	60.0	494	9	ADY14498	Ady14498	PRO	polyp	219	18	60.0	720	5	AAO14594	Aao14594	Lamini-n-r	
147	18	60.0	502	7	ADC35086	Adc35086	Human	bire	220	18	60.0	720	9	ADM47628	Adm47628	C. albica	
148	18	60.0	513	8	ADP31681	Adp31681	Human	sec	221	18	60.0	722	8	ADP98843	Adp98843		
149	18	60.0	514	8	ADP31122	Adp31122	Human	sec	222	18	60.0	750	8	ADP31131	Adp31131	Human	sec
150	18	60.0	516	7	ABM85365	Abm85365	Human	pro	223	18	60.0	756	8	ADP31039	Adp31039	Human	sec
151	18	60.0	516	8	ADP30871	Adp30871	Human	sec	224	18	60.0	757	8	ADP30925	Adp30925	Human	sec
152	18	60.0	516	8	ADP31418	Adp31418	Human	sec	225	18	60.0	768	8	ADP31149	Adp31149	Human	sec
153	18	60.0	531	8	ADP31539	Adp31539	Human	sec	226	18	60.0	768	8	ADP30912	Adp30912	Human	sec
154	18	60.0	531	8	ADP30594	Adp30594	Human	sec	227	18	60.0	768	8	ADP31362	Adp31362	Human	sec
155	18	60.0	531	8	ADP31540	Adp31540	Human	sec	228	18	60.0	768	8	ADP31363	Adp31363	Human	sec
156	18	60.0	537	8	ADP31696	Adp31696	Human	sec	229	18	60.0	771	8	ADP30907	Adp30907	Human	sec
157	18	60.0	537	8	ADP31557	Adp31557	Human	sec	230	18	60.0	771	8	ADP31469	Adp31469	Human	sec
158	18	60.0	538	8	ADY22481	Ady22481	Plant	full	231	18	60.0	771	8	ADP31244	Adp31244	Human	sec
159	18	60.0	541	2	ADH11359	Adh11359	Vertebrat		232	18	60.0	774	8	ADP30885	Adp30885	Human	sec
160	18	60.0	543	8	ADP30864	Adp30864	Human	sec	233	18	60.0	774	8	ADP30506	Adp30506	Human	sec
161	18	60.0	546	8	ADO19176	Ado19176	Human	PRO	234	18	60.0	774	8	ADP31225	Adp31225	Human	sec
162	18	60.0	546	8	ADP31449	Adp31449	Human	sec	235	18	60.0	779	8	ADP30915	Adp30915	Human	sec
163	18	60.0	546	8	ADP54223	Adp54223	Human	PRO	236	18	60.0	780	8	ADP30897	Adp30897	Human	sec
164	18	60.0	546	9	ADY14500	Ady14500	PRO	polyp	237	18	60.0	780	8	ADP31077	Adp31077	Human	sec
165	18	60.0	549	8	ADP31009	Adp31009	Human	sec	238	18	60.0	783	8	ADP31436	Adp31436	Human	sec
166	18	60.0	552	8	ADP30875	Adp30875	Human	sec	239	18	60.0	783	8	ADP30496	Adp30496	Human	sec
167	18	60.0	552	8	ADP31019	Adp31019	Human	sec	240	18	60.0	786	8	ADP30454	Adp30454	Human	sec
168	18	60.0	555	8	ADP31168	Adp31168	Human	sec	241	18	60.0	795	8	ADP31332	Adp31332	Human	sec
169	18	60.0	557	8	ADP31103	Adp31103	Human	sec	242	18	60.0	804	8	ADP31635	Adp31635	Human	sec
170	18	60.0	558	8	ADP31254	Adp31254	Human	sec	243	18	60.0	804	8	ADP31291	Adp31291	Human	sec

244	18	60.0	807	8	ADP31036	Adp31036	Human	sec	317	18	60.0	1092	8	ADP31153	Adp31153	Human	sec
245	18	60.0	810	8	ADP31307	Adp31307	Human	sec	318	18	60.0	1092	8	ADP31358	Adp31358	Human	sec
246	18	60.0	813	8	ADP30649	Adp30649	Human	sec	319	18	60.0	1107	8	ADP30482	Adp30482	Human	sec
247	18	60.0	813	8	ADP30561	Adp30561	Human	sec	320	18	60.0	1107	8	ADP30504	Adp30504	Human	sec
248	18	60.0	816	8	ADP31249	Adp31249	Human	sec	321	18	60.0	1110	8	ADP31430	Adp31430	Human	sec
249	18	60.0	816	8	ADP31581	Adp31581	Human	sec	322	18	60.0	1110	8	ADP31508	Adp31508	Human	sec
250	18	60.0	828	8	ADP31569	Adp31569	Human	sec	323	18	60.0	1116	8	ADP31203	Adp31203	Human	sec
251	18	60.0	831	8	ADP31333	Adp31333	Human	sec	324	18	60.0	1116	8	ADP31692	Adp31692	Human	sec
252	18	60.0	831	8	ADP31123	Adp31123	Human	sec	325	18	60.0	1116	8	ADP31202	Adp31202	Human	sec
253	18	60.0	831	8	ADP31179	Adp31179	Human	sec	326	18	60.0	1128	6	ADA15725	Ada15725	C. elegans	
254	18	60.0	843	3	AA444348	Aay44348	Human	hep	327	18	60.0	1133	8	ADP30931	Adp30931	Human	sec
255	18	60.0	843	3	ADP30663	Adp30663	Human	sec	328	18	60.0	1133	8	ADP30928	Adp30928	Human	sec
256	18	60.0	849	8	ADP31182	Adp31182	Human	sec	329	18	60.0	1133	8	ADP30933	Adp30933	Human	sec
257	18	60.0	861	8	ADP31020	Adp31020	Human	sec	330	18	60.0	1133	8	ADP30939	Adp30939	Human	sec
258	18	60.0	861	8	ADP31021	Adp31021	Human	sec	331	18	60.0	1134	8	ADP30647	Adp30647	Human	sec
259	18	60.0	865	8	ADP30566	Adp30566	Human	sec	332	18	60.0	1140	8	ADP31128	Adp31128	Human	sec
260	18	60.0	869	7	AB068000	Ab068000	Pseudomon		333	18	60.0	1140	8	ADP31130	Adp31130	Human	sec
261	18	60.0	870	8	ADP30646	Adp30646	Human	sec	334	18	60.0	1141	8	ADP30677	Adp30677	Human	sec
262	18	60.0	882	8	ADP31688	Adp31688	Human	sec	335	18	60.0	1142	8	ADP30929	Adp30929	Human	sec
263	18	60.0	885	8	ADP31198	Adp31198	Human	sec	336	18	60.0	1155	8	ADP31455	Adp31455	Human	sec
264	18	60.0	887	8	ADP30554	Adp30554	Human	sec	337	18	60.0	1165	8	ADP30879	Adp30879	Human	sec
265	18	60.0	887	8	ADP30548	Adp30548	Human	sec	338	18	60.0	1168	8	ADP31046	Adp31046	Human	sec
266	18	60.0	889	8	ADP31648	Adp31648	Human	sec	339	18	60.0	1170	8	ADP30919	Adp30919	Human	sec
267	18	60.0	890	8	ADP31059	Adp31059	Human	sec	340	18	60.0	1170	8	ADP30922	Adp30922	Human	sec
268	18	60.0	892	8	ADP31578	Adp31578	Human	sec	341	18	60.0	1173	8	ADP31155	Adp31155	Human	sec
269	18	60.0	897	8	ADP30914	Adp30914	Human	sec	342	18	60.0	1189	8	ADP31043	Adp31043	Human	sec
270	18	60.0	906	8	ADP31409	Adp31409	Human	sec	343	18	60.0	1192	8	ADP30964	Adp30964	Human	sec
271	18	60.0	918	8	ADP31459	Adp31459	Human	sec	344	18	60.0	1192	8	ADP31180	Adp31180	Human	sec
272	18	60.0	921	8	ADP31522	Adp31522	Human	sec	345	18	60.0	1197	8	ADP31034	Adp31034	Human	sec
273	18	60.0	925	5	AA014246	Aa014246	Human	pre	346	18	60.0	1199	8	ADP31044	Adp31044	Human	sec
274	18	60.0	930	8	ADP31444	Adp31444	Human	sec	347	18	60.0	1215	8	ADP30898	Adp30898	Human	sec
275	18	60.0	933	8	ADP31510	Adp31510	Human	sec	348	18	60.0	1215	8	ADP30900	Adp30900	Human	sec
276	18	60.0	936	8	ADP31597	Adp31597	Human	sec	349	18	60.0	1221	8	ADP31293	Adp31293	Human	sec
277	18	60.0	936	8	ADP31568	Adp31568	Human	sec	350	18	60.0	1224	8	ADP31426	Adp31426	Human	sec
278	18	60.0	939	8	ADP31542	Adp31542	Human	sec	351	18	60.0	1233	8	ADP30523	Adp30523	Human	sec
279	18	60.0	939	8	ADP31086	Adp31086	Human	sec	352	18	60.0	1239	8	ADP31297	Adp31297	Human	sec
280	18	60.0	939	8	ADP31541	Adp31541	Human	sec	353	18	60.0	1252	8	ADP30678	Adp30678	Human	sec
281	18	60.0	945	8	ADP31238	Adp31238	Human	sec	354	18	60.0	1260	8	ADP31533	Adp31533	Human	sec
282	18	60.0	945	8	ADP31237	Adp31237	Human	sec	355	18	60.0	1269	8	ADP31382	Adp31382	Human	sec
283	18	60.0	947	8	ADP30937	Adp30937	Human	sec	356	18	60.0	1269	8	ADP31381	Adp31381	Human	sec
284	18	60.0	948	8	ADP30586	Adp30586	Human	sec	357	18	60.0	1282	8	ADP31328	Adp31328	Human	sec
285	18	60.0	950	8	ADP31167	Adp31167	Human	sec	358	18	60.0	1288	8	ADP31114	Adp31114	Human	sec
286	18	60.0	954	8	ADP31015	Adp31015	Human	sec	359	18	60.0	1289	8	ADP30675	Adp30675	Human	sec
287	18	60.0	957	8	ADP31528	Adp31528	Human	sec	360	18	60.0	1302	8	ADP31695	Adp31695	Human	sec
288	18	60.0	960	8	ADP31471	Adp31471	Human	sec	361	18	60.0	1312	8	ADP30999	Adp30999	Human	sec
289	18	60.0	960	8	ADP31470	Adp31470	Human	sec	362	18	60.0	1312	8	ADP31382	Adp31382	Human	sec
290	18	60.0	966	8	ADP30745	Adp30745	Human	sec	363	18	60.0	1358	8	ADP30995	Adp30995	Human	sec
291	18	60.0	975	8	ADP30843	Adp30843	Human	sec	364	18	60.0	1362	8	ADP31181	Adp31181	Human	sec
292	18	60.0	990	8	ADP31553	Adp31553	Human	sec	365	18	60.0	1365	8	ADP31035	Adp31035	Human	sec
293	18	60.0	1002	8	ADG39639	Adg39639	Human	pan	366	18	60.0	1371	8	ADP31646	Adp31646	Human	sec
294	18	60.0	1002	8	ADP30866	Adp30866	Human	sec	367	18	60.0	1371	8	ADP30876	Adp30876	Human	sec
295	18	60.0	1010	8	ADP31296	Adp31296	Human	sec	368	18	60.0	1380	8	ADP31566	Adp31566	Human	sec
296	18	60.0	1017	6	AA026719	Aa026719	SR protei		369	18	60.0	1383	8	ADP31091	Adp31091	Human	sec
297	18	60.0	1017	6	ADP31268	Adp31268	Human	sec	370	18	60.0	1387	8	ADP30946	Adp30946	Human	sec
298	18	60.0	1023	8	ADP30994	Adp30994	Human	sec	371	18	60.0	1404	8	ADP31627	Adp31627	Human	sec
299	18	60.0	1030	8	ADP30913	Adp30913	Human	sec	372	18	60.0	1420	8	ADP30944	Adp30944	Human	sec
300	18	60.0	1033	8	ADP30984	Adp30984	Human	sec	373	18	60.0	1431	8	ADP31609	Adp31609	Human	sec
301	18	60.0	1035	8	ADP31552	Adp31552	Human	sec	374	18	60.0	1437	8	ADP31357	Adp31357	Human	sec
302	18	60.0	1038	8	ADP30860	Adp30860	Human	sec	375	18	60.0	1440	8	ADG34533	Adg34533	Glucocort	
303	18	60.0	1048	8	ADP31642	Adp31642	Human	sec	376	18	60.0	1454	8	ADP31177	Adp31177	Human	sec
304	18	60.0	1056	8	ADP31082	Adp31082	Human	sec	377	18	60.0	1456	8	ADP30923	Adp30923	Human	sec
305	18	60.0	1057	8	ADP31592	Adp31592	Human	sec	378	18	60.0	1464	8	ADP31040	Adp31040	Human	sec
306	18	60.0	1059	8	ADP31042	Adp31042	Human	sec	379	18	60.0	1464	8	ADP31437	Adp31437	Human	sec
307	18	60.0	1065	8	ADP30973	Adp30973	Human	sec	380	18	60.0	1470	8	ADP31390	Adp31390	Human	sec
308	18	60.0	1065	8	ADP31347	Adp31347	Human	sec	381	18	60.0	1472	8	ADP31611	Adp31611	Human	sec
309	18	60.0	1076	6	AB081145	Ab081145	Human	PRO	382	18	60.0	1480	8	ADP30557	Adp30557	Human	sec
310	18	60.0	1076	6	ABU66845	Abu66845	Human	PRO	383	18	60.0	1485	8	ADP31383	Adp31383	Human	sec
311	18	60.0	1081	2	AA424319	Aay424319	Mouse	dep	384	18	60.0	1485	8	ADP31384	Adp31384	Human	sec
312	18	60.0	1083	8	ADP31073	Adp31073	Human	sec	385	18	60.0	1488	8	ADP31385	Adp31385	Human	sec
313	18	60.0	1086	8	ADP31175	Adp31175	Human	sec	386	18	60.0	1488	8	ADP31386	Adp31386	Human	sec
314	18	60.0	1086	8	ADP31447	Adp31447	Human	sec	387	18	60.0	1498	8	ADP30684	Adp30684	Human	sec
315	18	60.0	1086	8	ADP31629	Adp31629	Human	sec	388	18	60.0	1506	8	ADP30596	Adp30596	Human	sec
316	18	60.0	1090	8	ADQ10187	Adq10187	Human	pol	389	18	60.0	1510	7	ADG30698	Adg30698	Xanthomon	

390	18	60.0	1518	8	ADP31532	Adp31532	Human	sec	463	18	60.0	1894	6	ABO05229	AbO05229	Novel	hum
391	18	60.0	1530	8	ADP31536	Adp31536	Human	sec	464	18	60.0	1917	6	ADA15719	AdA15719	C. elegans	
392	18	60.0	1539	8	ADP31201	Adp31201	Human	sec	465	18	60.0	1933	8	ADP30889	Adp30889	Human	sec
393	18	60.0	1539	8	ADP31200	Adp31200	Human	sec	466	18	60.0	1933	8	ADP30902	Adp30902	Human	sec
394	18	60.0	1550	8	ADP30567	Adp30567	Human	sec	467	18	60.0	2001	8	ADP31644	Adp31644	Human	sec
395	18	60.0	1584	8	ADP31405	Adp31405	Human	sec	468	18	60.0	2020	8	ADP31056	Adp31056	Human	sec
396	18	60.0	1587	8	ADP30581	Adp30581	Human	sec	469	18	60.0	2020	8	ADP30511	Adp30511	Human	sec
397	18	60.0	1614	8	ADP31529	Adp31529	Human	sec	470	18	60.0	2020	8	ADP30512	Adp30512	Human	sec
398	18	60.0	1617	8	ADP30660	Adp30660	Human	sec	471	18	60.0	2020	8	ADP30509	Adp30509	Human	sec
399	18	60.0	1623	8	ADP30552	Adp30552	Human	sec	472	18	60.0	2027	8	ADP31058	Adp31058	Human	sec
400	18	60.0	1629	8	ADP30945	Adp30945	Human	sec	473	18	60.0	2044	8	ADP30510	Adp30510	Human	sec
401	18	60.0	1647	8	ADP30670	Adp30670	Human	sec	474	18	60.0	2088	8	ADP31178	Adp31178	Human	sec
402	18	60.0	1647	8	ADP31052	Adp31052	Human	sec	475	18	60.0	2091	8	ADP31088	Adp31088	Human	sec
403	18	60.0	1652	6	ADA15715	AdA15715	C. elegans		476	18	60.0	2123	8	ADP30657	Adp30657	Human	sec
404	18	60.0	1665	8	ADP31187	Adp31187	Human	sec	477	18	60.0	2123	8	ADP30564	Adp30564	Human	sec
405	18	60.0	1719	8	ADP31137	Adp31137	Human	sec	478	18	60.0	2127	8	ADP31327	Adp31327	Human	sec
406	18	60.0	1725	8	ADP30654	Adp30654	Human	sec	479	18	60.0	2148	8	ADP30974	Adp30974	Human	sec
407	18	60.0	1743	6	ABU88255	Abu88255	Human	sec	480	18	60.0	2187	8	ADP30882	Adp30882	Human	sec
408	18	60.0	1743	6	ABU90134	Abu90134	Novel	hum	481	18	60.0	2260	8	ADP30687	Adp30687	Human	sec
409	18	60.0	1743	6	ABU96436	Abu96436	Novel	hum	482	18	60.0	2272	8	ADP30669	Adp30669	Human	sec
410	18	60.0	1743	6	ABU99045	Abu99045	Novel	hum	483	18	60.0	2304	8	ADP31252	Adp31252	Human	sec
411	18	60.0	1743	6	ABU98260	Abu98260	Novel	hum	484	18	60.0	2307	8	ADP31394	Adp31394	Human	sec
412	18	60.0	1743	6	ABU91966	Abu91966	Novel	hum	485	18	60.0	2349	8	ADP30959	Adp30959	Human	sec
413	18	60.0	1743	6	ABU85270	Abu85270	Novel	hum	486	18	60.0	2358	8	ADP31690	Adp31690	Human	sec
414	18	60.0	1743	6	ABO00409	AbO00409	Novel	hum	487	18	60.0	2382	8	ADP31341	Adp31341	Human	sec
415	18	60.0	1743	6	ABU88960	Abu88960	Novel	hum	488	18	60.0	2391	8	ADP31366	Adp31366	Human	sec
416	18	60.0	1743	6	ABO06456	AbO06456	Novel	hum	489	18	60.0	2415	8	ADP31023	Adp31023	Human	sec
417	18	60.0	1743	6	ABU95516	Abu95516	Novel	hum	490	18	60.0	2418	8	ADP31105	Adp31105	Human	sec
418	18	60.0	1743	6	ABU95206	Abu95206	Novel	hum	491	18	60.0	2454	8	ADP30469	Adp30469	Human	sec
419	18	60.0	1743	6	ABU90754	Abu90754	Novel	hum	492	18	60.0	2468	8	ADP30947	Adp30947	Human	sec
420	18	60.0	1743	6	ABU93916	Abu93916	Novel	hum	493	18	60.0	2469	8	ADP31031	Adp31031	Human	sec
421	18	60.0	1743	6	ABU86190	Abu86190	Novel	hum	494	18	60.0	2484	8	ADP66690	Adp66690	Human	mis
422	18	60.0	1743	6	ABU82045	Abu82045	Novel	hum	495	18	60.0	2508	6	ADA15721	AdA15721	C. elegans	
423	18	60.0	1743	6	ABU07906	Abu07906	Novel	hum	496	18	60.0	2514	8	ADP30727	Adp30727	Human	sec
424	18	60.0	1743	6	ABU94226	Abu94226	Novel	hum	497	18	60.0	2535	8	ADP31146	Adp31146	Human	sec
425	18	60.0	1743	6	ABO00099	AbO00099	Novel	hum	498	18	60.0	2544	6	ADA15717	AdA15717	C. elegans	
426	18	60.0	1743	6	ABU87110	Abu87110	Novel	hum	499	18	60.0	2601	6	ADA15723	AdA15723	C. elegans	
427	18	60.0	1743	6	ABU91351	Abu91351	Novel	hum	500	18	60.0	2611	8	ADP31027	Adp31027	Human	sec
428	18	60.0	1743	6	ABU90444	Abu90444	Novel	hum	501	18	60.0	2616	8	ADP31253	Adp31253	Human	sec
429	18	60.0	1743	6	ABU97035	Abu97035	Novel	hum	502	18	60.0	2616	9	ABE49675	AbE49675	N. mening	
430	18	60.0	1743	6	ABO05231	AbO05231	Novel	hum	503	18	60.0	2664	8	ADP31546	Adp31546	Human	sec
431	18	60.0	1755	8	ADP31446	Adp31446	Human	sec	504	18	60.0	2688	8	ADN11593	Adn11593	Human	CD9
432	18	60.0	1762	9	ADP62715	Adp62715	Human	alp	505	18	60.0	2808	8	ADP31258	Adp31258	Human	sec
433	18	60.0	1782	8	ADP31351	Adp31351	Human	sec	506	18	60.0	2832	8	ADP31121	Adp31121	Human	sec
434	18	60.0	1782	8	ADP31270	Adp31270	Human	sec	507	18	60.0	2835	8	ADP30572	Adp30572	Human	sec
435	18	60.0	1789	8	ADP30562	Adp30562	Human	sec	508	18	60.0	2843	8	ADP31120	Adp31120	Human	sec
436	18	60.0	1803	8	ADP30562	Adp30562	Human	sec	509	18	60.0	2980	9	AEA36049	AEa36049	Maize	Str
437	18	60.0	1827	8	ADP31170	Adp31170	Human	sec	510	18	60.0	3046	8	ADQ10419	Adq10419	Human	po1
438	18	60.0	1827	8	ADP30642	Adp30642	Human	sec	511	18	60.0	3070	8	ADP30710	Adp30710	Human	sec
439	18	60.0	1848	8	ADP31372	Adp31372	Human	sec	512	18	60.0	3070	8	ADP31544	Adp31544	Human	sec
440	18	60.0	1894	6	ABU88253	Abu88253	Novel	hum	513	18	60.0	3144	8	ADP31545	Adp31545	Human	sec
441	18	60.0	1894	6	ABU90132	Abu90132	Novel	hum	514	18	60.0	3201	8	ADP31545	Adp31545	Human	sec
442	18	60.0	1894	6	ABU96434	Abu96434	Novel	hum	515	18	60.0	3316	8	ADP31116	Adp31116	Human	sec
443	18	60.0	1894	6	ABU99043	Abu99043	Novel	hum	516	18	60.0	3339	8	ADP31219	Adp31219	Human	sec
444	18	60.0	1894	6	ABU98258	Abu98258	Novel	hum	517	18	60.0	3390	8	ADP31148	Adp31148	Human	sec
445	18	60.0	1894	6	ABU91964	Abu91964	Novel	hum	518	18	60.0	3398	9	ABE87634	AbE87634	Human	ino
446	18	60.0	1894	6	ABU85268	Abu85268	Novel	hum	519	18	60.0	3411	8	ADP30667	Adp30667	Human	sec
447	18	60.0	1894	6	ABO00407	AbO00407	Novel	hum	520	18	60.0	3447	8	ADP31112	Adp31112	Human	sec
448	18	60.0	1894	6	ABU88958	Abu88958	Novel	hum	521	18	60.0	3465	8	ADP31234	Adp31234	Human	sec
449	18	60.0	1894	6	ABO06454	AbO06454	Novel	hum	522	18	60.0	3477	8	ADP30704	Adp30704	Human	sec
450	18	60.0	1894	6	ABU95514	Abu95514	Novel	hum	523	18	60.0	3579	8	ADP31098	Adp31098	Human	sec
451	18	60.0	1894	6	ABU95204	Abu95204	Novel	hum	524	18	60.0	3585	8	ADP31117	Adp31117	Human	sec
452	18	60.0	1894	6	ABU90752	Abu90752	Novel	hum	525	18	60.0	4440	6	ABU88256	Abu88256	Novel	hum
453	18	60.0	1894	6	ABU93914	Abu93914	Novel	hum	526	18	60.0	4440	6	ABU90135	Abu90135	Novel	hum
454	18	60.0	1894	6	ABU86188	Abu86188	Novel	hum	527	18	60.0	4440	6	ABU96437	Abu96437	Novel	hum
455	18	60.0	1894	6	ABU82043	Abu82043	Novel	hum	528	18	60.0	4440	6	ABU99046	Abu99046	Novel	hum
456	18	60.0	1894	6	ABU07904	Abu07904	Novel	hum	529	18	60.0	4440	6	ABU98261	Abu98261	Novel	hum
457	18	60.0	1894	6	ABU94224	Abu94224	Novel	hum	530	18	60.0	4440	6	ABU91967	Abu91967	Novel	hum
458	18	60.0	1894	6	ABO00097	AbO00097	Novel	hum	531	18	60.0	4440	6	ABU85271	Abu85271	Novel	hum
459	18	60.0	1894	6	ABU87108	Abu87108	Novel	hum	532	18	60.0	4440	6	ABO00410	AbO00410	Novel	hum
460	18	60.0	1894	6	ABU91349	Abu91349	Novel	hum	533	18	60.0	4440	6	ABU88961	Abu88961	Novel	hum
461	18	60.0	1894	6	ABU90442	Abu90442	Novel	hum	534	18	60.0	4440	6	ABO06457	AbO06457	Novel	hum
462	18	60.0	1894	6	ABU97033	Abu97033	Novel	hum	535	18	60.0	4440	6	ABU95517	Abu95517	Novel	hum

536	18	60.0	4440	6	ABU95207	Novel hum	609	17	56.7	53	6	ABM54353	Abm54353	Proprionib
537	18	60.0	4440	6	ABU90755	Novel hum	610	17	56.7	53	6	ABM53330	Abm53330	Proprionib
538	18	60.0	4440	6	ABU93917	Novel hum	611	17	56.7	54	9	AEA30185	Aea30185	Pertussis
539	18	60.0	4440	6	ABU86191	Novel hum	612	17	56.7	58	4	AAU54758	Aau54758	Proprionib
540	18	60.0	4440	6	ABU82046	Novel hum	613	17	56.7	58	6	ABM51277	Abm51277	Proprionib
541	18	60.0	4440	6	ABU07907	Novel hum	614	17	56.7	58	9	ADZ11761	Adz11761	E. coli D
542	18	60.0	4440	6	ABU94227	Novel hum	615	17	56.7	58	9	AEA30433	Aea30433	Pertussis
543	18	60.0	4440	6	ABO00100	Novel hum	616	17	56.7	58	9	AEA30308	Aea30308	Pertussis
544	18	60.0	4440	6	ABU87111	Novel hum	617	17	56.7	59	4	AAU51154	Aau51154	Proprionib
545	18	60.0	4440	6	ABU91352	Novel hum	618	17	56.7	59	6	ABM47673	Abm47673	Proprionib
546	18	60.0	4440	6	ABU90445	Novel hum	619	17	56.7	60	9	AEA30430	Aea30430	Pertussis
547	18	60.0	4440	6	ABU97036	Novel hum	620	17	56.7	60	9	AEA30259	Aea30259	Pertussis
548	18	60.0	4440	6	ABO05232	Novel hum	621	17	56.7	60	9	AEA30493	Aea30493	Pertussis
549	18	60.0	4683	8	ADP31260	Human sec	622	17	56.7	60	9	AEA30376	Aea30376	Pertussis
550	18	60.0	4752	8	ADP30585	Human sec	623	17	56.7	60	9	AEA30494	Aea30494	Pertussis
551	18	60.0	4752	8	ADP30651	Human sec	624	17	56.7	60	9	AEA30431	Aea30431	Pertussis
552	18	60.0	4848	8	ADP31259	Human sec	625	17	56.7	60	9	AEA30432	Aea30432	Pertussis
553	18	60.0	5304	8	ADP30706	Human sec	626	17	56.7	62	7	ADJ38115	Adj38115	Human par
554	18	60.0	5397	8	ADP31068	Human sec	627	17	56.7	62	7	ABO69433	AbO69433	Pseudomon
555	18	60.0	5514	8	ADP31186	Human sec	628	17	56.7	64	2	AAV12328	Aav12328	Human 5'
556	18	60.0	5514	8	ADP31591	Human sec	629	17	56.7	64	2	AAW67942	Aaw67942	Fragment
557	18	60.0	5820	8	ADP31118	Human sec	630	17	56.7	64	4	AAW85119	Aaw85119	Human imm
558	18	60.0	6465	8	ADP30705	Human sec	631	17	56.7	64	5	ABP69086	Abp69086	Human pol
559	18	60.0	6729	8	ADP31600	Human sec	632	17	56.7	66	4	ABG13638	Abg13638	Novel hum
560	18	60.0	7285	6	ABJ38280	pAMG21 -RA	633	17	56.7	67	7	ABM86677	Abm86677	Rice abio
561	18	60.0	7339	6	AAO16358	Human tra	634	17	56.7	68	7	ABP09377	Abp09377	Human ORF
562	18	60.0	8973	8	ADP31119	Human sec	635	17	56.7	68	7	ABO78891	AbO78891	Pseudomon
563	18	60.0	8976	8	ADP31425	Human sec	636	17	56.7	70	4	AAU53345	Aau53345	Proprionib
564	18	60.0	9195	8	ADP31494	Human sec	637	17	56.7	70	6	ABM49864	Abm49864	Proprionib
565	18	60.0	10944	8	ADP31310	Human sec	638	17	56.7	71	2	AAW83934	Aaw83934	Human sec
566	18	60.0	11328	8	ADP31310	Human sec	639	17	56.7	71	3	AAAB34715	Aaab34715	Human sec
567	17	56.7	8	5	ABBA47073	Desmocoll	640	17	56.7	73	7	ABO82698	AbO82698	Pseudomon
568	17	56.7	8	5	ABBA6610	Desmocoll	641	17	56.7	73	8	ADP29540	Adp29540	Human sec
569	17	56.7	8	5	ADZ50990	Amino aci	642	17	56.7	74	4	AAU59878	Aau59878	Proprionib
570	17	56.7	8	9	ADZ50974	Amino aci	643	17	56.7	74	6	ABM56397	Abm56397	Proprionib
571	17	56.7	10	4	AAAG95975	Human com	644	17	56.7	77	9	ADV76639	Adv76639	Human TER
572	17	56.7	10	4	AAAG95917	Human com	645	17	56.7	78	8	ADX73420	Adx73420	Plant ful
573	17	56.7	20	2	AAV23344	Peptide 8	646	17	56.7	78	9	ABM94485	Abm94485	M. xanthu
574	17	56.7	20	3	AAI19697	O-Mechylc	647	17	56.7	79	4	AAAI19029	Aai19029	Proprionib
575	17	56.7	20	4	AAE05831	Liquidamb	648	17	56.7	79	4	ABR38235	AbR38235	Proprionib
576	17	56.7	20	9	ADZ14872	Liquidamb	649	17	56.7	79	4	AAW31668	Aaw31668	Proprionib
577	17	56.7	22	4	AAW33651	Peptide #	650	17	56.7	79	4	ABR23413	AbR23413	Protein #
578	17	56.7	22	4	AAW73453	Human bon	651	17	56.7	79	4	ABG53088	AbG53088	Human liv
579	17	56.7	22	4	AAW60779	Human bra	652	17	56.7	79	5	ABG41186	AbG41186	Human pep
580	17	56.7	22	4	ABG55176	Human liv	653	17	56.7	80	3	AAG11389	Aag11389	Proprionib
581	17	56.7	22	5	ABG43312	Human pep	654	17	56.7	83	4	AAU64692	Aau64692	Proprionib
582	17	56.7	24	8	ADP73753	Influenza	655	17	56.7	83	6	ABM61211	Abm61211	Proprionib
583	17	56.7	28	9	AAAG98776	Human cel	656	17	56.7	84	3	AAAG60503	Aag60503	Arabidops
584	17	56.7	28	9	ADZ47463	Human HAS	657	17	56.7	84	3	AAAG55303	Aag55303	Arabidops
585	17	56.7	34	5	AAU87022	Adenoviru	658	17	56.7	84	4	AAU61337	Aau61337	Proprionib
586	17	56.7	35	6	ABU61319	Human A d	659	17	56.7	84	6	ABM57856	Abm57856	Proprionib
587	17	56.7	35	8	ABU61408	Low densi	660	17	56.7	85	9	ABR41187	AbR41187	L. pneumo
588	17	56.7	35	8	ADP21541	Human LDL	661	17	56.7	86	4	AAU51874	Aau51874	Proprionib
589	17	56.7	37	4	AAO09823	Human pol	662	17	56.7	86	6	ABM48393	Abm48393	Proprionib
590	17	56.7	39	9	ABR43285	Human LRP	663	17	56.7	87	4	AAU63689	Aau63689	Proprionib
591	17	56.7	39	9	AAU59877	Proprionib	664	17	56.7	87	6	ABM60208	Abm60208	Proprionib
592	17	56.7	50	6	ABM56396	Proprionib	665	17	56.7	87	7	ADP17451	Adp17451	Mouse IL-
593	17	56.7	51	4	ABR39045	Peptide #	666	17	56.7	88	4	AAU21754	Aau21754	Novel hum
594	17	56.7	51	4	AAW32534	Peptide #	667	17	56.7	88	4	AAU45755	Aau45755	Proprionib
595	17	56.7	51	4	AAW72275	Human bon	668	17	56.7	88	4	AAU20119	Aau20119	Human dna
596	17	56.7	51	4	AAW59697	Human bra	669	17	56.7	88	5	ABG91368	Abg91368	Novel hum
597	17	56.7	51	4	ABG53961	Human liv	670	17	56.7	88	6	ABM42274	Abm42274	Proprionib
598	17	56.7	51	5	ABG42090	Human pep	671	17	56.7	88	7	ADC46395	Adc46395	Human neo
599	17	56.7	53	4	AAW19878	Peptide #	672	17	56.7	88	7	ABO69228	AbO69228	Pseudomon
600	17	56.7	53	4	ABR39811	Peptide #	673	17	56.7	90	8	ADP31655	Adp31655	Human sec
601	17	56.7	53	4	AAW33401	Peptide #	674	17	56.7	91	2	AAV12296	Aav12296	Human 5'
602	17	56.7	53	4	ABR24424	Protein #	675	17	56.7	92	5	ABO09944	AbO09944	Guinea pi
603	17	56.7	53	4	AAW73191	Human bon	676	17	56.7	92	5	ABR09943	AbR09943	Human int
604	17	56.7	53	4	AAU56801	Proprionib	677	17	56.7	92	7	AAE39396	Aae39396	Guinea pi
605	17	56.7	53	4	AAU57834	Proprionib	678	17	56.7	92	7	AAE39395	Aae39395	Human int
606	17	56.7	53	4	AAW60535	Human bra	679	17	56.7	92	9	AEA35800	Aea35800	Partial i
607	17	56.7	53	4	ABG54907	Human liv	680	17	56.7	92	9	AEA35802	Aea35802	Partial i
608	17	56.7	53	5	ABG43038	Human pep	681	17	56.7	93	4	AAU56711	Aau56711	Proprionib

682	17	56.7	93	6	ABM53230	Abm53230	Propionib	755	17	56.7	134	6	ABM55230	Abm55230	Propionib
683	17	56.7	96	4	AAU52513	Aau52513	Propionib	756	17	56.7	134	7	ABO76893	AbO76893	Pseudomon
684	17	56.7	96	6	ABM49032	Abm49032	Propionib	757	17	56.7	135	3	AAV75297	Aay75297	Neisseria
685	17	56.7	97	4	AAU66144	Aau66144	Propionib	758	17	56.7	135	5	ABP08959	Abp08959	Human ORF
686	17	56.7	97	5	ABP09938	Abp09938	Human ORF	759	17	56.7	135	5	ABO74109	AbO74109	Pseudomon
687	17	56.7	97	6	ABM62663	Abm62663	Propionib	760	17	56.7	135	8	ADP31638	Adp31638	Human sec
688	17	56.7	97	7	ABO79445	AbO79445	Pseudomon	761	17	56.7	135	8	ADP31638	Adp31638	Human sec
689	17	56.7	97	9	ABE37882	Aeb37882	L. pneumo	762	17	56.7	136	8	ADU07492	AdU07492	Plant full
690	17	56.7	99	4	AAU49282	Aau49282	Propionib	763	17	56.7	136	3	AAV75295	Aay75295	Neisseria
691	17	56.7	99	5	ABP43926	Abp43926	NICE-1 pr	764	17	56.7	136	7	ADM57545	Adm57545	Human fls
692	17	56.7	99	6	ABM45781	Abm45781	Propionib	765	17	56.7	136	7	ABO73988	AbO73988	Pseudomon
693	17	56.7	99	8	ADN04970	Adn04970	Antipsori	766	17	56.7	136	7	ABM66687	Abm66687	Rice abio
694	17	56.7	102	7	ABO79727	AbO79727	Pseudomon	767	17	56.7	137	3	AAV65413	Aay65413	Human 5'
695	17	56.7	105	7	ADBE6267	Adbe6267	Rat Prote	768	17	56.7	137	8	ADU72977	AdU72977	Signal pe
696	17	56.7	106	4	AAAG64065	Aaag64065	Human ant	769	17	56.7	137	9	ADZ73968	Adz73968	Human com
697	17	56.7	106	4	AAAG64068	Aaag64068	Rat anter	770	17	56.7	140	7	ADF13949	Adf13949	Human end
698	17	56.7	106	5	ABM15333	Abm15333	Human nar	771	17	56.7	140	7	ABO82130	AbO82130	Pseudomon
699	17	56.7	106	5	AAU96154	Aau96154	Human nar	772	17	56.7	141	7	ABO77083	AbO77083	Pseudomon
700	17	56.7	106	6	ADA57463	Ada57463	Human sec	773	17	56.7	143	5	AAU96140	Aau96140	Human BRP
701	17	56.7	106	6	ADA41343	Ada41343	Human sec	774	17	56.7	144	8	ABO79215	AbO79215	Pseudomon
702	17	56.7	106	6	ABR48065	AbR48065	Human sec	775	17	56.7	144	8	ADP31474	Adp31474	Human pro
703	17	56.7	106	6	ABP72423	Abp72423	Human gly	776	17	56.7	144	8	ADP31474	Adp31474	Human sec
704	17	56.7	106	8	ADF72484	Adf72484	Mature hu	777	17	56.7	144	8	ADP31474	Adp31474	Human end
705	17	56.7	106	9	ADM86161	Adm86161	Human Zlu	778	17	56.7	148	7	ADP13954	Adp13954	Human end
706	17	56.7	107	7	ADP13950	Adp13950	Human end	779	17	56.7	148	7	ABO74263	AbO74263	Pseudomon
707	17	56.7	107	8	ADP30723	Adp30723	Human sec	780	17	56.7	149	4	AAU68926	Aau68926	Human pro
708	17	56.7	111	4	ABG03591	Abg03591	Novel hum	781	17	56.7	149	7	ADP31059	Adp31059	Human dia
709	17	56.7	112	7	ABO77337	AbO77337	Novel hum	782	17	56.7	149	7	ADP13246	Adp13246	Protease
710	17	56.7	113	4	AAU50472	Aau50472	Pseudomon	783	17	56.7	149	8	ADH78455	Adh78455	Human pro
711	17	56.7	113	6	ABM46991	Abm46991	Propionib	784	17	56.7	150	8	ADK66524	Adk66524	Plant full
712	17	56.7	113	8	ADP31617	Adp31617	Human sec	785	17	56.7	154	7	ABO74102	AbO74102	Pseudomon
713	17	56.7	113	8	ADP31613	Adp31613	Human sec	786	17	56.7	154	8	ABO60522	AbO60522	Human gen
714	17	56.7	113	3	AAE25477	Aae25477	Pinus rad	787	17	56.7	156	7	ABO80146	AbO80146	Pseudomon
715	17	56.7	117	7	ADBE5158	Adbe5158	Human pro	788	17	56.7	156	8	ADP30759	Adp30759	Human sec
716	17	56.7	117	7	ABO68458	AbO68458	Pseudomon	789	17	56.7	157	8	ADK78826	Adk78826	Plant full
717	17	56.7	118	6	ABU99121	Abu99121	Novel hum	790	17	56.7	162	3	AAV95932	Aay95932	Porcine a
718	17	56.7	118	8	ADM93807	Adm93807	Human NOV	791	17	56.7	162	4	AAU42696	Aau42696	Propionib
719	17	56.7	120	8	ADB64538	Adb64538	Human pro	792	17	56.7	162	6	ABM39215	Abm39215	Propionib
720	17	56.7	120	8	ADP30556	Adp30556	Human sec	793	17	56.7	163	7	ABO80365	AbO80365	Pseudomon
721	17	56.7	122	4	AAU42386	Aau42386	Propionib	794	17	56.7	163	7	ABO68705	AbO68705	Pseudomon
722	17	56.7	122	6	ABM38905	Abm38905	Propionib	795	17	56.7	164	7	ABO78378	AbO78378	Pseudomon
723	17	56.7	123	7	ABO66168	AbO66168	Klebsiell	796	17	56.7	165	4	AAU58075	Aau58075	Propionib
724	17	56.7	124	4	AAU51146	Aau51146	Propionib	797	17	56.7	165	6	ABM54594	Abm54594	Propionib
725	17	56.7	124	6	ABM47665	Abm47665	Propionib	798	17	56.7	165	8	ADP31286	Adp31286	Human sec
726	17	56.7	127	8	ADP31622	Adp31622	Human sec	799	17	56.7	165	8	ADP31515	Adp31515	Human sec
727	17	56.7	129	4	AAAG64067	Aaag64067	Rat anter	800	17	56.7	165	8	ADP31174	Adp31174	Human sec
728	17	56.7	129	5	AAU96157	Aau96157	Human BRP	801	17	56.7	165	8	ADP31516	Adp31516	Human sec
729	17	56.7	129	5	AAU96129	Aau96129	Human bet	802	17	56.7	168	8	ADP31621	Adp31621	Human sec
730	17	56.7	129	6	ABU97054	Abu97054	Recombina	803	17	56.7	168	8	ADP30797	Adp30797	Human sec
731	17	56.7	130	4	AAAG63211	Aaag63211	Maitro acI	804	17	56.7	168	8	ADP30769	Adp30769	Human sec
732	17	56.7	130	4	AAAG64064	Aaag64064	Human ant	805	17	56.7	168	8	ADP30771	Adp30771	Human sec
733	17	56.7	130	4	AAU41366	Aau41366	Propionib	806	17	56.7	171	8	ADP30599	Adp30599	Human sec
734	17	56.7	130	5	AAE09440	Aae09440	Human sbg	807	17	56.7	171	8	ADP30599	Adp30599	Human sec
735	17	56.7	130	5	AAU97612	Aau97612	Human OGH	808	17	56.7	175	3	AAAG54788	Aaag54788	Arabidops
736	17	56.7	130	5	AAU96153	Aau96153	Human BRP	809	17	56.7	175	7	ABO80411	AbO80411	Pseudomon
737	17	56.7	130	5	AAU10368	Aau10368	Mouse bet	810	17	56.7	175	7	ABO69379	AbO69379	Pseudomon
738	17	56.7	130	5	AAU10366	Aau10366	Human bet	811	17	56.7	177	7	ABO70036	AbO70036	Pseudomon
739	17	56.7	130	6	ABG74215	Abg74215	Mouse gly	812	17	56.7	178	7	ABO74750	AbO74750	Pseudomon
740	17	56.7	130	6	ABG74206	Abg74206	Human gly	813	17	56.7	178	7	ABO73284	AbO73284	Pseudomon
741	17	56.7	130	6	ABM37885	Abm37885	Propionib	814	17	56.7	179	7	ABO75778	AbO75778	Pseudomon
742	17	56.7	130	6	ABP72432	Abp72432	Human gly	815	17	56.7	180	8	ADP30820	Adp30820	Human sec
743	17	56.7	130	7	ADJ38121	Adj38121	Human OGH	816	17	56.7	180	8	ADP30825	Adp30825	Human sec
744	17	56.7	130	7	ABO81191	AbO81191	Pseudomon	817	17	56.7	180	8	ADP30828	Adp30828	Human sec
745	17	56.7	130	8	ADF72483	Adf72483	Human GPH	818	17	56.7	180	8	ADP30821	Adp30821	Human sec
746	17	56.7	130	8	ADM6155	Adm6155	Human Zlu	819	17	56.7	180	8	ADP30826	Adp30826	Human sec
747	17	56.7	130	9	ADM6175	Adm6175	Martine zI	820	17	56.7	180	8	ADP30827	Adp30827	Human sec
748	17	56.7	130	9	AE812349	Aeb12349	Human BRP	821	17	56.7	182	7	ABO78099	AbO78099	Pseudomon
749	17	56.7	131	5	AAU96155	Aau96155	Human BRP	822	17	56.7	182	7	ABO71666	AbO71666	Pseudomon
750	17	56.7	133	4	AAU49190	Aau49190	Propionib	823	17	56.7	182	8	ADN99530	Adn99530	Novel hum
751	17	56.7	133	4	ABG04654	Abg04654	Novel hum	824	17	56.7	184	7	ABO77341	AbO77341	Pseudomon
752	17	56.7	133	6	ABM45709	Abm45709	Propionib	825	17	56.7	185	8	ADP31109	Adp31109	Human sec
753	17	56.7	133	7	ABO82412	AbO82412	Pseudomon	826	17	56.7	186	4	ABG08184	AbG08184	Novel hum
754	17	56.7	134	4	AAU58711	Aau58711	Propionib	827	17	56.7	187	7	ABO80088	AbO80088	Pseudomon

828	17	56.7	190	6	ABM65810	Abm65810	Propionib	901	17	56.7	249	3	AA660495	AA660495	Arabidops
829	17	56.7	191	7	ABO76619	ABO76619	Pseudomon	902	17	56.7	249	7	ABO83546	ABO83546	Abob3546
830	17	56.7	192	8	ADP31335	ADP31335	Human sec	903	17	56.7	250	6	ABP56993	E. tenell	Abp56993
831	17	56.7	192	8	ADP30575	ADP30575	Human sec	904	17	56.7	250	7	ABO82689	Pseudomon	Abj32689
832	17	56.7	198	4	AAU47307	AAU47307	Propionib	905	17	56.7	251	7	ADJ92161	Human hai	Adj92161
833	17	56.7	198	6	ABM43826	ABM43826	Propionib	906	17	56.7	251	7	ABO73872	Human hai	Abj3872
834	17	56.7	198	8	ADP30811	ADP30811	Human sec	907	17	56.7	251	9	ABM91192	M. xanthu	Abm91192
835	17	56.7	198	8	ADP30811	ADP30811	Human sec	908	17	56.7	254	8	ADP30739	Human sec	Adp30739
836	17	56.7	198	8	ADP30477	ADP30477	Human sec	909	17	56.7	254	8	ADP31396	Human hai	Adp31396
837	17	56.7	198	8	ADP30481	ADP30481	Human sec	910	17	56.7	255	7	ADJ92145	Human hai	Adj92145
838	17	56.7	200	7	ABO71047	ABO71047	Human sec	911	17	56.7	255	8	ADP30658	Human sec	Adp30658
839	17	56.7	201	7	ABO60877	ABO60877	Klebsiell	912	17	56.7	256	4	ABG08337	Novel hum	Abg08337
840	17	56.7	203	7	ADB80007	ADB80007	Mycobacte	913	17	56.7	256	6	ABU97138	Recombin	Abu97138
841	17	56.7	204	8	ABO71404	ABO71404	Human sec	914	17	56.7	256	7	ABO75228	Pseudomon	AbO75228
842	17	56.7	206	7	ABO74022	ABO74022	Pseudomon	915	17	56.7	256	8	ADN99494	Novel hum	Adn99494
843	17	56.7	208	4	AAU29506	AAU29506	Human G p	916	17	56.7	257	5	ABP64573	Human ORF	Abp64573
844	17	56.7	208	5	ABG60794	ABG60794	Novel G p	917	17	56.7	258	8	ADP31084	Human sec	Adp31084
845	17	56.7	208	7	ABO80531	ABO80531	Pseudomon	918	17	56.7	259	7	ABM88128	Rice abio	Abm88128
846	17	56.7	209	7	ABO77021	ABO77021	Pseudomon	919	17	56.7	261	6	AAO27174	Arabidops	AAO27174
847	17	56.7	210	8	ADP31563	ADP31563	Human sec	920	17	56.7	263	7	AA660494	Arabidops	AA660494
848	17	56.7	210	8	ADP31562	ADP31562	Human sec	921	17	56.7	263	7	ABO77554	Pseudomon	AbO77554
849	17	56.7	212	7	ABO67532	ABO67532	Klebsiell	922	17	56.7	263	8	ABM81215	Tumour -48	Abm81215
850	17	56.7	215	7	ADBO8255	ADBO8255	Novel pro	923	17	56.7	264	8	ADP30788	Human sec	Adp30788
851	17	56.7	215	7	ABO72061	ABO72061	Human sec	924	17	56.7	264	8	ADP31456	Human sec	Adp31456
852	17	56.7	215	9	ABM93776	ABM93776	M. xanthu	925	17	56.7	264	8	ADP31412	Human sec	Adp31412
853	17	56.7	217	4	AAJ97773	AAJ97773	I. scapul	926	17	56.7	266	8	ADP31427	Human sec	Adp31427
854	17	56.7	218	7	ABO81158	ABO81158	Pseudomon	927	17	56.7	266	8	ADP31427	Human sec	Adp31427
855	17	56.7	218	8	ADY23534	ADY23534	Plant ful	928	17	56.7	267	4	ABG27981	Novel hum	Abg27981
856	17	56.7	219	8	ADP31171	ADP31171	Human sec	929	17	56.7	267	8	ADP31483	Human sec	Adp31483
857	17	56.7	219	8	ADY06385	ADY06385	Plant ful	930	17	56.7	270	8	ADP31217	Human sec	Adp31217
858	17	56.7	221	7	ADJ92147	ADJ92147	Human hai	931	17	56.7	270	8	ADP31321	Human sec	Adp31321
859	17	56.7	221	7	ADM87372	ADM87372	Human pro	932	17	56.7	270	8	ADP31564	Human sec	Adp31564
860	17	56.7	222	5	ABM78572	ABM78572	Rat ribon	933	17	56.7	270	9	ABE12365	Fusion pr	Abi2365
861	17	56.7	222	5	ABM78573	ABM78573	Mouse rib	934	17	56.7	271	6	ABU36614	Protein e	Abu36614
862	17	56.7	223	3	AA803862	AA803862	Human neu	935	17	56.7	271	6	ABU34590	Protein e	Abu34590
863	17	56.7	224	2	AAW22018	AAW22018	Pig myoge	936	17	56.7	271	6	ADJ92151	Human hai	Adj92151
864	17	56.7	224	9	ADW95939	ADW95939	Human myo	937	17	56.7	272	6	ABU35898	Protein e	Abu35898
865	17	56.7	224	9	ADX15805	ADX15805	Human myo	938	17	56.7	272	8	ADP12954	Protein e	Adp12954
866	17	56.7	225	5	ABM78570	ABM78570	Human ryo	939	17	56.7	273	8	ADP31236	Human sec	Adp31236
867	17	56.7	225	7	ABM88876	ABM88876	Rice abio	940	17	56.7	273	9	ABM94710	Human dia	Abm94710
868	17	56.7	225	8	ADP31501	ADP31501	Human sec	941	17	56.7	274	4	AAU19418	Human dia	AAU19418
869	17	56.7	225	8	ADP30777	ADP30777	Human sec	942	17	56.7	274	7	ABO72003	Pseudomon	AbO72003
870	17	56.7	227	7	ABO75458	ABO75458	Pseudomon	943	17	56.7	274	8	ADP31679	Human sec	Adp31679
871	17	56.7	228	8	ADP30921	ADP30921	Human sec	944	17	56.7	275	6	ABU34080	Protein e	Abu34080
872	17	56.7	231	5	ABP64726	ABP64726	Human pro	945	17	56.7	276	7	ABO75093	Pseudomon	AbO75093
873	17	56.7	233	7	ABO75933	ABO75933	Pseudomon	946	17	56.7	276	8	ADP31319	Human sec	Adp31319
874	17	56.7	235	7	ABO69554	ABO69554	Pseudomon	947	17	56.7	276	8	ADP30584	Human sec	Adp30584
875	17	56.7	237	8	ADP31322	ADP31322	Human sec	948	17	56.7	278	8	ADP31314	Human sec	Adp31314
876	17	56.7	238	9	ABE12367	ABE12367	Fusion pr	949	17	56.7	279	8	ADP30846	Human sec	Adp30846
877	17	56.7	239	3	AAAB21310	AAAB21310	Human zym	950	17	56.7	279	8	ADP31523	Human sec	Adp31523
878	17	56.7	239	8	ADY13038	ADY13038	Plant ful	951	17	56.7	279	8	ADP31503	Human sec	Adp31503
879	17	56.7	240	2	AAJ22205	AAJ22205	Biorthythm	952	17	56.7	279	8	ADP31671	Human sec	Adp31671
880	17	56.7	240	2	ADP30737	ADP30737	Human sec	953	17	56.7	279	8	ADP30605	Human sec	Adp30605
881	17	56.7	244	2	AAK44532	AAK44532	Hyme App-	954	17	56.7	279	8	ADP30844	Human sec	Adp30844
882	17	56.7	244	2	AAW22985	AAW22985	Human ser	955	17	56.7	282	7	ADJ92143	Human hai	Adj92143
883	17	56.7	244	2	AAW51006	AAW51006	Protease	956	17	56.7	285	7	ABO73521	Pseudomon	AbO73521
884	17	56.7	244	3	AAAB21323	AAAB21323	Human zym	957	17	56.7	288	8	ADP31463	Human sec	Adp31463
885	17	56.7	244	5	ABG96357	ABG96357	Human ova	958	17	56.7	291	7	ABO68749	Pseudomon	AbO68749
886	17	56.7	244	6	AAE37572	AAE37572	Human 204	959	17	56.7	293	6	ADA55092	Human pro	Ada55092
887	17	56.7	244	7	ADB80567	ADB80567	Ovarian c	960	17	56.7	293	8	ADH09610	Human hos	Adh09610
888	17	56.7	244	7	ADN39212	ADN39212	Cancer/an	961	17	56.7	294	8	ADP31076	Human sec	Adp31076
889	17	56.7	244	8	ADJ39734	ADJ39734	Human pro	962	17	56.7	294	8	ADP31473	Human sec	Adp31473
890	17	56.7	244	8	ADJ37158	ADJ37158	Human pro	963	17	56.7	297	8	ADP31192	Human sec	Adp31192
891	17	56.7	244	8	ADNO4074	ADNO4074	Antipsoi	964	17	56.7	298	2	AAW43397	Human int	AAW43397
892	17	56.7	244	8	ADN29289	ADN29289	Human kal	965	17	56.7	302	3	AAO07023	Arabidops	AAO07023
893	17	56.7	244	8	ADQ89076	ADQ89076	Human uro	966	17	56.7	304	8	ADP31659	Human sec	Adp31659
894	17	56.7	244	8	ADR72624	ADR72624	Human xen	967	17	56.7	306	7	ADFO4021	Bacterial	Adfo4021
895	17	56.7	244	8	ADR72876	ADR72876	Human ova	968	17	56.7	307	7	ADCO1287	Enteroha	Adco1287
896	17	56.7	244	9	ADY67594	ADY67594	Human kal	969	17	56.7	309	8	ADI42146	Plant tra	Adi42146
897	17	56.7	244	9	ADZ51357	ADZ51357	Amino aci	970	17	56.7	310	4	ABBS8993	Drosophi	Abbs8993
898	17	56.7	245	7	ABO78304	ABO78304	Pseudomon	971	17	56.7	311	4	ABM67853	Human hai	Abm67853
899	17	56.7	246	8	ADP30619	ADP30619	Human sec	972	17	56.7	311	4	ABG08405	Novel hum	Abg08405
900	17	56.7	246	9	ADY18495	ADY18495	PRO polyP	973	17	56.7	312	1	AAp50079	T-cell an	AAp50079

```
974 17 56.7 312 1 AaP60471 AaP60471 Portion o
975 17 56.7 312 2 AaR53145 AaR53145 T-cell an
976 17 56.7 312 3 AaB25555 AaB25555 Pinus rad
977 17 56.7 312 4 AaG08404 AaG08404 Novel hum
978 17 56.7 312 8 ADP31505 ADP31505 Human sec
979 17 56.7 312 8 ADP30476 ADP30476 Human sec
980 17 56.7 313 7 ABM85966 ABM85966 Rice abio
981 17 56.7 315 5 AaE13348 AaE13348 Human TST
982 17 56.7 315 6 ABR58684 ABR58684 Human can
983 17 56.7 315 8 ADP31685 ADP31685 Human tas
984 17 56.7 316 3 AaY44862 AaY44862 Human sec
985 17 56.7 316 9 ADY18921 ADY18921 PRO polyP
986 17 56.7 317 8 ADP31615 ADP31615 Human sec
987 17 56.7 318 7 AB078573 AB078573 Pseudomon
988 17 56.7 318 8 ADP31066 ADP31066 Human sec
989 17 56.7 320 4 AaG91631 AaG91631 C glutamI
990 17 56.7 325 3 AaB51703 AaB51703 Gene 34 h
991 17 56.7 326 3 AaG24798 AaG24798 Arabidops
992 17 56.7 336 8 ADP31320 ADP31320 Human sec
993 17 56.7 341 7 AB077462 AB077462 Pseudomon
994 17 56.7 343 7 ADP13953 ADP13953 Human end
995 17 56.7 345 3 AaG07022 AaG07022 Arabidops
996 17 56.7 345 3 AaG37538 AaG37538 Arabidops
997 17 56.7 345 8 ADP31204 ADP31204 Human sec
998 17 56.7 346 7 AB081210 AB081210 Pseudomon
999 17 56.7 346 7 ABM89803 ABM89803 Rice abio
1000 17 56.7 346 8 ADP31355 ADP31355 Human sec
```

ALIGNMENTS

RESULT 1

ADQ13022 standard; peptide; 9 AA.

```
ADQ13022:
07-OCT-2004 (first entry)
Hepatitis B virus MHC class I restricted T-cell stimulating peptide 821.
HBV, MHC class I restricted T-cell stimulating peptide; surface peptide;
core peptide; polymerase peptide; immunogenic composition; vaccine;
cytotoxic T-lymphocyte response; CTL response.
Hepatitis B virus.
WO2004058807-A2.
15-JUL-2004.
09-DEC-2003; 2003WO-EP013948.
24-DEC-2002; 2002EP-00447276.
(ALGO-) ALGONOMICS NV.
Lasters I, Desmet J, Stegmann T, Castelein B;
WPI, 2004-525861/50.
New peptide comprising a major histocompatibility complex class I
restricted T-cell stimulating epitope of the hepatitis B virus (HBV)
surface, core and/or polymerase polypeptide, for preparation of an HBV
immunogenic composition.
Claim 5; SEQ ID NO 821; 108bp; English.
The invention comprises major histocompatibility complex (MHC) class I
restricted T-cell stimulating epitopes of the hepatitis B virus (HBV)
surface, core and polymerase proteins. The peptides of the invention are
useful for generating an HBV immunogenic composition (e.g. vaccine) that
```

CC induces a cytotoxic T-lymphocyte (CTL) response. The present amino acid
CC sequence represents an MHC class I restricted T-cell stimulating HBV
CC peptide of the invention.

SQ Sequence 9 AA;

Query Match 60.0%; Score 18; DB 8; Length 9;
Best Local Similarity 25.0%; Pred. No. 1.4e+06;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 1 CASSSSSC 8

RESULT 2

ADQ13021 standard; peptide; 9 AA.

```
ADQ13021:
07-OCT-2004 (first entry)
Hepatitis B virus MHC class I restricted T-cell stimulating peptide 820.
HBV, MHC class I restricted T-cell stimulating peptide; surface peptide;
core peptide; polymerase peptide; immunogenic composition; vaccine;
cytotoxic T-lymphocyte response; CTL response.
Hepatitis B virus.
WO2004058807-A2.
15-JUL-2004.
09-DEC-2003; 2003WO-EP013948.
24-DEC-2002; 2002EP-00447276.
(ALGO-) ALGONOMICS NV.
Lasters I, Desmet J, Stegmann T, Castelein B;
WPI, 2004-525861/50.
New peptide comprising a major histocompatibility complex class I
restricted T-cell stimulating epitope of the hepatitis B virus (HBV)
surface, core and/or polymerase polypeptide, for preparation of an HBV
immunogenic composition.
Example 7; SEQ ID NO 820; 108bp; English.
The invention comprises major histocompatibility complex (MHC) class I
restricted T-cell stimulating epitopes of the hepatitis B virus (HBV)
surface, core and polymerase proteins. The peptides of the invention are
useful for generating an HBV immunogenic composition (e.g. vaccine) that
induces a cytotoxic T-lymphocyte (CTL) response. The present amino acid
sequence represents an MHC class I restricted T-cell stimulating HBV
peptide of the invention.
```

SQ Sequence 9 AA;

Query Match 60.0%; Score 18; DB 8; Length 9;
Best Local Similarity 25.0%; Pred. No. 1.4e+06;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 2 CASSSSSC 9

RESULT 3

AAW65943

```

ID AAW65943 standard; peptide; 11 AA.
XX
AC AAW65943;
XX
DT 12-NOV-1998 (first entry)
XX
DE Molecule VI.
XX
KM Neurotrophin; cyclic; bicyclic; disulphide; nerve growth factor; NGF;
XX BDNF; NT-3; conformation; promoter.
XX
OS Synthetic.
OS Mammalia.
XX
PN CA2205045-A.
XX
PD 12-MAY-1998.
XX
PF 09-MAY-1997; 97CA-02205045.
XX
PR 12-NOV-1996; 96CA-02190296.
XX
PA (TOOH ) UNIV QUEBENS KINGSTON.
XX
PI Riopelle RJ, Dory MI, Shamovsky IL, Weaver DF, Ross GM;
XX
DR WPI; 1998-457650/40.
XX
PT Inhibition of neurotrophin activity - using factor that interferes with
XX sub-unit interaction.
XX
PS Disclosure; Page 13; 103pp; English.
XX
CC The invention relates to a method for reducing the biological activity of
CC a multimeric protein having at least 2 promoters. It comprises (a)
CC providing a factor that interacts with at least one portion of at least
CC one of the promoters which associates with a portion of the other
CC promoter in the absence of the factor; and (b) mixing the factor with the
CC multimeric protein so that the factor interacts with the portion(s) and
CC disrupts association of at least a portion of the promoters. The method
CC and compounds are useful for inhibiting neurotrophin-mediated activities
CC selected from neurotrophin receptor binding, neuron survival, neurite
CC outgrowth and epileptic effects. The present sequence is shown in the
CC specification
XX
SQ Sequence 11 AA;
XX
Query Match 60.0%; Score 18; DB 2; Length 11;
Best Local Similarity 25.0%; Pred. No. 2.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
DB 1 CAAAAAAC 8

```

RESULT 4
ID AAW65944 standard; peptide; 11 AA.
XX
AC AAW65944;
XX
DT 12-NOV-1998 (first entry)
XX
DE Molecule XV.
XX
KM Neurotrophin; cyclic; bicyclic; disulphide; nerve growth factor; NGF;
XX BDNF; NT-3; conformation; promoter.
XX
OS Synthetic.
OS Mammalia.
XX
PN CA2205045-A.

```

XX
PD 12-MAY-1998.
XX
PF 09-MAY-1997; 97CA-02205045.
XX
PR 12-NOV-1996; 96CA-02190296.
XX
PA (TOOH ) UNIV QUEBENS KINGSTON.
XX
PI Riopelle RJ, Dory MI, Shamovsky IL, Weaver DF, Ross GM;
XX
DR WPI; 1998-457650/40.
XX
PT Inhibition of neurotrophin activity - using factor that interferes with
XX sub-unit interaction.
XX
PS Disclosure; Page 13; 103pp; English.
XX
CC The invention relates to a method for reducing the biological activity of
CC a multimeric protein having at least 2 promoters. It comprises (a)
CC providing a factor that interacts with at least one portion of at least
CC one of the promoters which associates with a portion of the other
CC promoter in the absence of the factor; and (b) mixing the factor with the
CC multimeric protein so that the factor interacts with the portion(s) and
CC disrupts association of at least a portion of the promoters. The method
CC and compounds are useful for inhibiting neurotrophin-mediated activities
CC selected from neurotrophin receptor binding, neuron survival, neurite
CC outgrowth and epileptic effects. The present sequence is shown in the
CC specification
XX
SQ Sequence 11 AA;
XX
Query Match 60.0%; Score 18; DB 2; Length 11;
Best Local Similarity 25.0%; Pred. No. 2.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
DB 4 CAAAAAAC 11

```

RESULT 5
ID ADV23688 standard; peptide; 15 AA.
XX
AC ADV23688;
XX
DT 10-MAR-2005 (first entry)
XX
DE HBV immunogenic peptide #489.
XX
KM Vaccine; virucide; antigen; autoimmune disease; infection;
KM immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;
KM breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;
KM pancreas tumor; stomach tumor; bladder tumor; kidney tumor;
KM Hodgkin's lymphoma.
XX
OS Hepatitis B virus.
XX
PN WO2004108753-A1.
XX
PD 16-DEC-2004.
XX
PF 10-JUN-2004; 2004WO-AU000775.
XX
PR 10-JUN-2003; 2003AU-00902875.
XX
PR 25-MAR-2004; 2004AU-00901589.
XX
PA (UYME) UNIV MELBOURNE.
XX
PI Kent SJ;
XX
DR WPI; 2005-031657/03.

XX Use of at least one set of peptides in the preparation of a medicament
 PT for modulating an immune response, and for treating cancer or yeast,
 PT viral, bacterial, protozoal and mycoplasma infections.

PS Disclosure; SEQ ID NO 2108; 645bp; English.

XX The invention relates to the use of at least one set of peptides in the
 CC preparation of a medicament for modulating an immune response, where
 CC individual peptides of a respective set comprise different portions of an
 CC amino acid sequence corresponding to a single polypeptide of interest and
 CC display partial sequence identity or similarity to at least one other
 CC peptide of the same set of peptides (i.e. they are overlapping). Also
 CC included are an antigen-presenting cell which has been contacted with the
 CC peptides above and thus presents the peptides, a population of such
 CC antigen-presenting cells, a process for producing antigen-presenting
 CC cells for modulating an immune response to a polypeptide of interest, a
 CC method for producing antigen-specific lymphocytes, a composition
 CC comprising at least one set of the peptides (and a carrier and/or
 CC diluent), a method for modulating an immune response to a polypeptide of
 CC interest comprising administering to a patient in need at least one set
 CC of the peptides, a method for treatment and/or prophylaxis of a disease
 CC or condition associated with the presence of a polypeptide of interest
 CC and a composition of matter for modulating an immune response in a
 CC subject to a target antigen. The polypeptide of interest is also a
 CC disease- or condition-associated polypeptide that is a polypeptide
 CC produced by a pathogenic organism or a cancer, and produced by a
 CC pathogenic organism selected from yeast, viruses, bacteria, helminths,
 CC protozoans and mycoplasmas. The disease- or condition-associated
 CC polypeptide is produced by a cancer selected from melanoma, lung cancer,
 CC breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic
 CC cancer, stomach cancer, bladder cancer, kidney cancer, post transplant
 CC lymphoproliferative disease (PTLD) or Hodgkin's lymphoma. The uncultured
 CC antigen-presenting cells or their precursors are useful in the
 CC preparation of a medicament for the treatment of a disease or condition
 CC in a subject, which disease or condition is associated with the presence
 CC of aberrant expression of a target antigen, where the antigen-presenting
 CC cells or their precursors have not been subjected to activating
 CC conditions but have been contacted with an antigen that corresponds to
 CC the target antigen to express a processed or modified form of the antigen
 CC for presentation to the subject's immune system. The present sequence is
 CC one of a set of overlapping immunogenic peptides derived from a Hepatitis
 CC B virus protein.

XX Sequence 15 AA;

XX Query Match 60.0%; Score 18; DB 9; Length 15;

XX Best Local Similarity 25.0%; Pred. No. 2.4e+02; Indels 0; Gaps 0;

XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

DB 4 CASSSSSC 11

RESULT 6

ADV23687 ADV23687 standard; peptide; 15 AA.

XX ADV23687;

XX 10-MAR-2005 (first entry)

DE HBV immunogenic peptide #488.

XX Vaccine; virucide; antigen; autoimmune disease; infection;

KW immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;

KW breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;

KW pancreas tumor; stomach tumor; bladder tumor; kidney tumor;

XX hodgkin's lymphoma.

XX Hepatitis B virus.

PN WO2004108753-A1.

XX 16-DEC-2004.

XX 10-JUN-2004; 2004WO-AU000775.

XX 10-JUN-2003; 2003AU-00902875.

XX 25-MAR-2004; 2004AU-00901589.

XX (UYME) UNIV MELBOURNE.

PI Kent SJ;

DR WPI; 2005-031657/03.

XX Use of at least one set of peptides in the preparation of a medicament
 PT for modulating an immune response, and for treating cancer or yeast,
 PT viral, bacterial, protozoal and mycoplasma infections.

PS Disclosure; SEQ ID NO 2107; 645bp; English.

XX The invention relates to the use of at least one set of peptides in the
 CC preparation of a medicament for modulating an immune response, where
 CC individual peptides of a respective set comprise different portions of an
 CC amino acid sequence corresponding to a single polypeptide of interest and
 CC display partial sequence identity or similarity to at least one other
 CC peptide of the same set of peptides (i.e. they are overlapping). Also
 CC included are an antigen-presenting cell which has been contacted with the
 CC peptides above and thus presents the peptides, a population of such
 CC antigen-presenting cells, a process for producing antigen-presenting
 CC cells for modulating an immune response to a polypeptide of interest, a
 CC method for producing antigen-specific lymphocytes, a composition
 CC comprising at least one set of the peptides (and a carrier and/or
 CC diluent), a method for modulating an immune response to a polypeptide of
 CC interest comprising administering to a patient in need at least one set
 CC of the peptides, a method for treatment and/or prophylaxis of a disease
 CC or condition associated with the presence of a polypeptide of interest
 CC and a composition of matter for modulating an immune response in a
 CC subject to a target antigen. The polypeptide of interest is also a
 CC disease- or condition-associated polypeptide that is a polypeptide
 CC produced by a pathogenic organism or a cancer, and produced by a
 CC pathogenic organism selected from yeast, viruses, bacteria, helminths,
 CC protozoans and mycoplasmas. The disease- or condition-associated
 CC polypeptide is produced by a cancer selected from melanoma, lung cancer,
 CC breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic
 CC cancer, stomach cancer, bladder cancer, kidney cancer, post transplant
 CC lymphoproliferative disease (PTLD) or Hodgkin's lymphoma. The uncultured
 CC antigen-presenting cells or their precursors are useful in the
 CC preparation of a medicament for the treatment of a disease or condition
 CC in a subject, which disease or condition is associated with the presence
 CC of aberrant expression of a target antigen, where the antigen-presenting
 CC cells or their precursors have not been subjected to activating
 CC conditions but have been contacted with an antigen that corresponds to
 CC the target antigen to express a processed or modified form of the antigen
 CC for presentation to the subject's immune system. The present sequence is
 CC one of a set of overlapping immunogenic peptides derived from a Hepatitis
 CC B virus protein.

XX Sequence 15 AA;

XX Query Match 60.0%; Score 18; DB 9; Length 15;

XX Best Local Similarity 25.0%; Pred. No. 2.4e+02; Indels 0; Gaps 0;

XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

DB 8 CASSSSSC 15

RESULT 7

ABBA42273 ABB42273 standard; peptide; 24 AA.

XX

```

AC ABB42273;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #9779 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 27; SEQ ID NO 34908; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 24 AA;

```

```

Query Match      60.0%; Score 18; DB 4; Length 24;
Best Local Similarity 25.0%; Pred. No. 2.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 4 CXXXXXXC 11
   |
DB 12 CTTSTTTC 19

```

```

RESULT 8
ID AAM36081 standard; protein; 24 AA.
XX
AC AAM36081;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #10118 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.

```

```

XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-48897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 36350; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see A131315-A157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 24 AA;

```

```

Query Match      60.0%; Score 18; DB 4; Length 24;
Best Local Similarity 25.0%; Pred. No. 2.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 4 CXXXXXXC 11
   |
DB 12 CTTSTTTC 19

```

```

RESULT 9
ID AAM75973 standard; protein; 24 AA.
XX
AC AAM75973;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 36279.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX

```

DR WPI; 2001-488900/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX Example 4; SEQ ID NO 36279; 658bp + Sequence Listing; English.
PS
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
SQ Sequence 24 AA;
Query Match 60.0%; Score 18; DB 4; Length 24;
Best Local Similarity 25.0%; Pred. No. 2.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
DB 12 CTTSTTTC 19
RESULT 10
AAM63159 standard; protein; 24 AA.
ID AAM63159;
AC AAM63159;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35264.
XX
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
OS Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
XX Example 4; SEQ ID NO 35264; 650bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX

SQ Sequence 24 AA;
Query Match 60.0%; Score 18; DB 4; Length 24;
Best Local Similarity 25.0%; Pred. No. 2.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
DB 12 CTTSTTTC 19
RESULT 11
ABG57702 standard; peptide; 24 AA.
ID ABG57702;
AC ABG57702;
XX
XX 25-FEB-2003 (first entry)
XX
XX Human liver peptide, SEQ ID No 36350.
XX
XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX
XX Homo sapiens.
XX
XX WO200157273-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000664.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488998/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
XX Claim 27; SEQ ID NO 36350; 658bp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 24 AA;
SQ
Query Match 60.0%; Score 18; DB 4; Length 24;
Best Local Similarity 25.0%; Pred. No. 2.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11

Db 12 CTTSTTTC 19

RESULT 12

ABBB68695 standard; protein; 69 AA.

ABBB68695;

26-MAR-2002 (first entry)

Drosophila melanogaster polypeptide SEQ ID NO 32877.

Drosophila; developmental biology; cell signalling; insecticide;

pharmaceutical.

Drosophila melanogaster.

MO200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001WO-US0092331.

23-MAR-2000; 2000US-0191637P.

11-JUL-2000; 2000US-00614150.

(PEKE) PE CORP NY.

Venter JC, Adams M, Li FMD, Myers EW;

WPI; 2001-656860/75.

N-PSDB; ABL12798.

New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.

Disclosure; SEQ ID NO 32877; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB16176-AB130511), expressed DNA sequences (AB101840-AB16175) and the encoded proteins (AB57737-AB872072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pct_sequences

Sequence 69 AA;

Query Match Best Local Similarity 60.0%; Score 18; DB 4; Length 69;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
|
9 CSSAASSC 16

RESULT 13

ADP30715 standard; protein; 69 AA.

ADP30715;

12-AUG-2004 (first entry)

Human secreted protein SEQ ID #1482.

KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

PN MO2004035732-A2.

29-APR-2004.

28-AUG-2003; 2003WO-US026780.

29-AUG-2002; 2002US-0406576P.

29-AUG-2002; 2002US-0406579P.

29-AUG-2002; 2002US-0406585P.

29-AUG-2002; 2002US-0406588P.

29-AUG-2002; 2002US-0406608P.

29-AUG-2002; 2002US-0406611P.

29-AUG-2002; 2002US-0406612P.

29-AUG-2002; 2002US-0406616P.

29-AUG-2002; 2002US-0406642P.

29-AUG-2002; 2002US-0406646P.

29-AUG-2002; 2002US-0406653P.

29-AUG-2002; 2002US-0406655P.

17-SEP-2002; 2002US-0410946P.

17-SEP-2002; 2002US-0410947P.

17-SEP-2002; 2002US-0410948P.

17-SEP-2002; 2002US-0410949P.

17-SEP-2002; 2002US-0410953P.

17-SEP-2002; 2002US-0410957P.

17-SEP-2002; 2002US-0410958P.

17-SEP-2002; 2002US-0410959P.

17-SEP-2002; 2002US-0410960P.

17-SEP-2002; 2002US-0410961P.

17-SEP-2002; 2002US-0410962P.

17-SEP-2002; 2002US-0411019P.

17-SEP-2002; 2002US-0411022P.

17-SEP-2002; 2002US-0411023P.

17-SEP-2002; 2002US-0411024P.

17-SEP-2002; 2002US-0411032P.

17-SEP-2002; 2002US-0411035P.

17-SEP-2002; 2002US-0411037P.

17-SEP-2002; 2002US-0411041P.

17-SEP-2002; 2002US-0411045P.

17-SEP-2002; 2002US-0411046P.

17-SEP-2002; 2002US-0411048P.

17-SEP-2002; 2002US-0411052P.

17-SEP-2002; 2002US-0411055P.

17-SEP-2002; 2002US-0411073P.

17-SEP-2002; 2002US-0411082P.

17-SEP-2002; 2002US-0411011P.

17-SEP-2002; 2002US-0411111P.

18-APR-2003; 2003US-0463700P.

18-APR-2003; 2003US-0463708P.

18-APR-2003; 2003US-0463716P.

18-APR-2003; 2003US-0463732P.

02-MAY-2003; 2003US-0467199P.

02-MAY-2003; 2003US-0467201P.

02-MAY-2003; 2003US-0467203P.

02-MAY-2003; 2003US-0467230P.

19-MAY-2003; 2003US-0471306P.

19-MAY-2003; 2003US-0471336P.

22-MAY-2003; 2003US-0472420P.

02-MAY-2003; 2003US-0472430P.

09-JUN-2003; 2003US-0476609P.

09-JUN-2003; 2003US-0476641P.

08-JUL-2003; 2003US-0485218P.

08-JUL-2003; 2003US-0485223P.

08-JUL-2003; 2003US-0485224P.

08-JUL-2003; 2003US-0485325P.

14-JUL-2003; 2003US-0486446P.

14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halanbeck RF, Huang MM, Kochakota S, Haishan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2713; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC anti-inflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
XX Sequence 69 AA;
SQ

Query Match 60.0%; Score 18; DB 8; Length 69;
Best Local Similarity 25.0%; Pred. No. 3.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
DB 18 CTTTAAAC 25

RESULT 14
ABB39058
ID ABB39058 standard; peptide; 91 AA.
XX
XX ABB39058;
XX
XX 04-FEB-2002 (first entry)
XX
XX Peptide #6564 encoded by human foetal liver single exon probe.
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
XX Homo sapiens.
XX WO200157277-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000669.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX

DR WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
XX Claim 27; SEQ ID NO 31693; 639pp + Sequence Listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probe may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPo at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 91 AA;
SQ

Query Match 60.0%; Score 18; DB 4; Length 91;
Best Local Similarity 25.0%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
DB 29 CSSSSSSC 36

RESULT 15
AAM32549
ID AAM32549 standard; protein; 91 AA.
XX
XX AAM32549;
XX
XX 17-OCT-2001 (first entry)
XX
XX Peptide #6586 encoded by probe for measuring placental gene expression.
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX
XX Homo sapiens.
XX WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
XX Claim 27; SEQ ID NO 32818; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP;
CC see AAI13135-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders

XX
SQ Sequence 91 AA;

Query Match 60.0%; Score 18; DB 4; Length 91;
Best Local Similarity 25.0%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 CXXXXXXC 11
|
DB 29 CXXXXSSC 36

RESULT 16

AAM72290
ID AAM72290 standard; protein; 91 AA.

XX
AC AAM72290;

XX
DT 06-NOV-2001 (first entry)

XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 32596.

XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
microarray; cancer; leukaemia; lymphoma; myeloma.

XX
OS Homo sapiens.

XX
PN WO200157276-A2.

XX
PD 09-AUG-2001.

XX
PF 30-JAN-2001; 2001WO-US000668.

XX
PR 04-FEB-2000; 2000US-0180312P.

XX
PR 26-MAY-2000; 2000US-0207456P.

XX
PR 30-JUN-2000; 2000US-00608408.

XX
PR 03-AUG-2000; 2000US-00632366.

XX
PR 21-SEP-2000; 2000US-0234687P.

XX
PR 27-SEP-2000; 2000US-0236359P.

XX
PR 04-OCT-2000; 2000GB-00024263.

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
DR WPI; 2001-488900/53.

XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX
PS gene expression in human bone marrow.

XX
PS Example 4; SEQ ID NO 32596; 658bp + Sequence Listing; English.

XX
CC The present invention provides a number of single exon nucleic acid

XX
CC probes which are derived from genomic sequences expressed in the human

XX
CC bone marrow. They can be used to measure gene expression in bone marrow

XX
CC samples, which may enable the improved diagnosis and treatment of cancers

XX
CC such as lymphoma, leukaemia and myeloma. The present sequence is a

XX
CC protein encoded by one of the probes of the invention

XX
SQ Sequence 91 AA;

Query Match 60.0%; Score 18; DB 4; Length 91;

Best Local Similarity 25.0%; Pred. No. 3.5e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 CXXXXXXC 11
|
DB 29 CXXXXSSC 36

RESULT 17

AAM59710
ID AAM59710 standard; protein; 91 AA.

XX
AC AAM59710;

XX
DT 05-NOV-2001 (first entry)

XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 31815.

XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;

XX
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

XX
OS Homo sapiens.

XX
PN WO200157275-A2.

XX
PD 09-AUG-2001.

XX
PF 30-JAN-2001; 2001WO-US000667.

XX
PR 04-FEB-2000; 2000US-0180312P.

XX
PR 26-MAY-2000; 2000US-0207456P.

XX
PR 30-JUN-2000; 2000US-00608408.

XX
PR 03-AUG-2000; 2000US-00632366.

XX
PR 21-SEP-2000; 2000US-0234687P.

XX
PR 27-SEP-2000; 2000US-0236359P.

XX
PR 04-OCT-2000; 2000GB-00024263.

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
DR WPI; 2001-483446/52.

XX
PT Single exon nucleic acid probes for analyzing gene expression in human

XX
PT brains.

XX
PS Example 4; SEQ ID NO 31815; 650bp + Sequence Listing; English.

XX
CC The present invention provides a number of single exon nucleic acid

XX
CC probes which are derived from genomic sequences expressed in the human

XX
CC brain. They can be used to measure gene expression in brain cell samples,

XX
CC which may enable the diagnosis and improved treatment of nervous system

XX
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

XX
CC epilepsy and cancers. The present sequence is a protein encoded by one of

XX
CC the probes of the invention

XX
SQ Sequence 91 AA;

Query Match 60.0%; Score 18; DB 4; Length 91;

Best Local Similarity 25.0%; Pred. No. 3.5e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 CXXXXXXC 11
|
DB 29 CXXXXSSC 36

RESULT 18

AAG53976
ID AAG53976 standard; peptide; 91 AA.

XX
AC AAG53976;

XX
DT 25-FEB-2003 (first entry)

XX
DE Human liver peptide, SEQ ID NO 32624.

XX
KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;

XX
KW hypercholesterolaemia; coronary heart disease.

XX
OS Homo sapiens.

XX
PN WO200157273-A2.

```

PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000664.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488698/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
XX Claim 27; SEQ ID NO 32624; 658bp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridizes at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (I) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABG47348-ABG5930 represent human
XX liver single exon encoded peptides of the invention. Note: The sequence
XX information for this patent does not appear in the printed specification
XX but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 91 AA;
XX
XX Query Match 60.0%; Score 18; DB 4; Length 91;
XX Best Local Similarity 25.0%; Pred. No. 3.5e+02;
XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
OY 4 CXXXXXXC 11
DB 29 CSSSSSSC 36
XX
RESULT 19
ABG42105
ID ABG42105 standard; peptide; 91 AA.
XX
XX ABG42105;
XX
XX 19-AUG-2002 (first entry)
XX
XX Human peptide encoded by genome-derived single exon probe SEQ ID 31770.
XX
XX Human; single exon probe; asthma; lung cancer; COPD; IHD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.
XX
XX Homo sapiens.
XX
XX WO200186003-A2.
XX

```

```

PD 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US000665.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX
XX Claim 27; SEQ ID NO 31770; 634p; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of probes
XX (I) the novel set of probes which hybridize at high stringency to a nucleic
XX acid expressed in the human lung; measuring gene expression in a sample
XX derived from human lung, comprising (a) contacting the array with a
XX collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of the
XX array; identifying exons in a eukaryotic genome, comprising (a)
XX algorithmically predicting at least one exon from genomic sequences of
XX the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarrays having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 12011 sequences, mentioned in the specification, or encoded by the
XX probes/open reading frames (ORF). The probes are used for gene expression
XX analysis, and for identifying exons in a gene, particularly using human
XX lung derived mRNA and for the study of lung diseases such as asthma, lung
XX cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
XX disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
XX tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
XX Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
XX histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
XX Karsenger syndrome, fibrocystic pulmonary dysplasia, primary ciliary
XX dyskinesia, pulmonary hypertension and hyaline membrane disease. The
XX present sequence is a peptide/protein encoded by a single exon probe of
XX the invention. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 91 AA;
XX
XX Query Match 60.0%; Score 18; DB 5; Length 91;
XX Best Local Similarity 25.0%; Pred. No. 3.5e+02;
XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
OY 4 CXXXXXXC 11
DB 29 CSSSSSSC 36
XX

```

```

RESULT 20
ADP30859
ID ADP30859 standard; protein; 93 AA.
XX
XX ADP30859;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1626.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX
XX 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX
XX 29-AUG-2002; 2002US-0406579P.
XX
XX 29-AUG-2002; 2002US-0406585P.
XX
XX 29-AUG-2002; 2002US-0406588P.
XX
XX 29-AUG-2002; 2002US-0406608P.
XX
XX 29-AUG-2002; 2002US-0406611P.
XX
XX 29-AUG-2002; 2002US-0406612P.
XX
XX 29-AUG-2002; 2002US-0406616P.
XX
XX 29-AUG-2002; 2002US-0406640P.
XX
XX 29-AUG-2002; 2002US-0406642P.
XX
XX 29-AUG-2002; 2002US-0406646P.
XX
XX 29-AUG-2002; 2002US-0406653P.
XX
XX 29-AUG-2002; 2002US-0406655P.
XX
XX 29-AUG-2002; 2002US-0406666P.
XX
XX 17-SEP-2002; 2002US-0410946P.
XX
XX 17-SEP-2002; 2002US-0410947P.
XX
XX 17-SEP-2002; 2002US-0410948P.
XX
XX 17-SEP-2002; 2002US-0410949P.
XX
XX 17-SEP-2002; 2002US-0410953P.
XX
XX 17-SEP-2002; 2002US-0410957P.
XX
XX 17-SEP-2002; 2002US-0410958P.
XX
XX 17-SEP-2002; 2002US-0410959P.
XX
XX 17-SEP-2002; 2002US-0410960P.
XX
XX 17-SEP-2002; 2002US-0410961P.
XX
XX 17-SEP-2002; 2002US-0410962P.
XX
XX 17-SEP-2002; 2002US-0411019P.
XX
XX 17-SEP-2002; 2002US-0411022P.
XX
XX 17-SEP-2002; 2002US-0411023P.
XX
XX 17-SEP-2002; 2002US-0411024P.
XX
XX 17-SEP-2002; 2002US-0411032P.
XX
XX 17-SEP-2002; 2002US-0411035P.
XX
XX 17-SEP-2002; 2002US-0411037P.
XX
XX 17-SEP-2002; 2002US-0411041P.
XX
XX 17-SEP-2002; 2002US-0411045P.
XX
XX 17-SEP-2002; 2002US-0411046P.
XX
XX 17-SEP-2002; 2002US-0411048P.
XX
XX 17-SEP-2002; 2002US-0411052P.
XX
XX 17-SEP-2002; 2002US-0411055P.
XX
XX 17-SEP-2002; 2002US-0411073P.
XX
XX 17-SEP-2002; 2002US-0411082P.
XX
XX 17-SEP-2002; 2002US-0411101P.
XX
XX 17-SEP-2002; 2002US-0411111P.
XX
XX 18-APR-2003; 2003US-0463700P.
XX
XX 18-APR-2003; 2003US-0463708P.
XX
XX 18-APR-2003; 2003US-0463716P.
XX
XX 18-APR-2003; 2003US-0463732P.
XX
XX 02-MAY-2003; 2003US-0467199P.
XX
XX 02-MAY-2003; 2003US-0467201P.
XX
XX 02-MAY-2003; 2003US-0467203P.
XX
XX 19-MAY-2003; 2003US-0467230P.
XX
XX 19-MAY-2003; 2003US-0471306P.
XX
XX 19-MAY-2003; 2003US-0471336P.

```

```

PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee R, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2857; 428pp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic.
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
XX Sequence 93 AA;
XX
XX Query Match 60.0%; Score 18; DB 8; Length 93;
XX Best Local Similarity 25.0%; Pred. No. 3.5e+02;
XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX QY 4 CXXXXXXC 11
XX
XX Db 85 CATATATAC 92
XX
XX RESULT 21
XX ABO83167
XX ID ABO83167 standard; protein; 99 AA.
XX
XX AC ABO83167;
XX
XX DT 29-JUL-2004 (first entry)
XX
XX DE Pseudomonas aeruginosa polypeptide #15342.
XX
XX DE Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX OS Pseudomonas aeruginosa.
XX
XX PN US6551795-B1.
XX
XX PD 22-APR-2003.
XX
XX PF 18-FEB-1999; 99US-00252991.
XX
XX XX 18-FEB-1998; 98US-0074788P.
XX
XX PR 27-JUL-1998; 98US-0094190P.

```

XX (GENO-) GENOME THERAPEUTICS CORP.
PA Rubenfield MJ, Nolling J, DeLoughery C, Bush D;
XX WPI: 2003-615309/58.
DR N-PSDB; ABD16738.
XX Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PR pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 31913; 455bp; English.
XX
XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-*P. aeruginosa* drugs, as templates for recombinant
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology. Sequences AB067826-
CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
SQ Sequence 99 AA:
Query Match 60.0%; Score 18; DB 7; Length 99;
Best Local Similarity 25.0%; Pred. No. 3.6e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
OY 4 CXXXXXXC 11
Db 24 CSSTRASAC 31
RESULT 22
ADP30709
ID ADP30709 standard; protein; 99 AA.
XX
AC ADP30709;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1476.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Halehan L, Linnemann T;
PI Plerce K, Wang Y, Wong JCP, Wu G, Zhang H;
XX WPI: 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 2707; 428bp; English.
XX

CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
SQ Sequence 99 AA;

Query Match 60.0%; Score 18; DB 8; Length 99;
Best Local Similarity 25.0%; Pred. No. 3.6e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 77 CAAATATC 84

RESULT 23

ID ABO74878 standard; protein; 107 AA.

XX ABO74878;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #7053.

KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

PN US651795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PA 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI; 2003-615309/58.

XX N-PSDB; ABD08449.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 23624; 455pp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using bioclip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX

SQ Sequence 107 AA;

Query Match 60.0%; Score 18; DB 7; Length 107;
Best Local Similarity 25.0%; Pred. No. 3.6e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 29 CSASSSAC 36

RESULT 24

ID ABO73394 standard; protein; 115 AA.

XX ABO73394;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #5569.

KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

PN US651795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PA 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI; 2003-615309/58.

XX N-PSDB; ABD06965.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 22140; 455pp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using bioclip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX

SQ Sequence 115 AA;

Query Match 60.0%; Score 18; DB 7; Length 115;
Best Local Similarity 25.0%; Pred. No. 3.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 83 CSAATSTC 90

RESULT 25
ADP04777
ID ADF04777 standard; protein; 122 AA.
XX
AC ADF04777;
XX
DT 12-FEB-2004 (first entry)
XX
DE Bacterial polypeptide #890.
XX
KM Proteus mirabilis infection; bacterial infection; antibacterial;
KW immunostimulant.
XX
OS Proteus mirabilis.
XX
PN US6605709-B1.
XX
PD 12-AUG-2003.
XX
PF 05-APR-2000; 2000US-00543681.
XX
PR 09-APR-1999; 99US-0128706P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL;
XX
DR MPI: 2003-895291/82.
XX
DR N-PSDB; ADP00605.
XX
PT New Proteus mirabilis polypeptides and polynucleotides, useful as
PT reagents for diagnosis of bacterial disease, as components of
PT antibacterial vaccines, as targets for antibacterial drugs, or as
PT biocontrol agents for plants.
XX
PS Disclosure; SEQ ID NO 5062; 870pp; English.
XX
CC The invention relates to new Proteus mirabilis polypeptides and
CC polynucleotides. The invention also relates to antibodies against the
CC polypeptides, methods for producing the polypeptides, a method of
CC generating vaccines for immunising an individual against P. mirabilis, a
CC method for evaluating a compound for the ability to bind a P. mirabilis
CC polypeptide and a method for screening test compounds for anti-bacterial
CC activity. The polypeptides and polynucleotides are useful as molecular
CC targets for diagnosing, preventing and treating pathological conditions
CC resulting from bacterial infection, as reagents for diagnosis of
CC bacterial diseases, as components of antibacterial vaccines, as targets
CC for antibacterial drugs or as bio-control agents for plants. This
CC sequence represents a Proteus mirabilis polypeptide of the invention.
XX
SQ Sequence 122 AA;

Query Match 60.0%; Score 18; DB 7; Length 122;
Best Local Similarity 25.0%; Pred. No. 3.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 CXXXXXXC 11
DB 99 CAASAC 106

RESULT 26
ADP30703
ID ADP30703 standard; protein; 135 AA.
XX
AC ADP30703;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1470.
XX

KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; Immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411077P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 19-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476099P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-048524P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halebeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2701; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
SQ Sequence 135 AA;

Query Match 60.0%; Score 18; DB 8; Length 135;
Best Local Similarity 25.0%; Pred. No. 3.8e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
DB 105 CTTTATC 112

RESULT 27
ABO75730
ID ABO75730 standard; protein; 144 AA.
XX
XX ABO75730;
AC
XX 29-JUL-2004 (first entry)
DT
XX
XX Pseudomonas aeruginosa polypeptide #7905.
DE
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
KM
XX Pseudomonas aeruginosa.
OS
XX
XX US6551795-B1.
PN
XX
XX 22-APR-2003.
PD
XX
XX 18-FEB-1999; 99US-00252991.
PF
XX
XX 18-FEB-1998; 98US-0074788P.
PR
XX 27-JUL-1998; 98US-0094190P.
PR
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI WPI; 2003-615309/58.
DR N-PSDB; ABD09301.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of

PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 24476; 455pp; English.
PS
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 144 AA;

Query Match 60.0%; Score 18; DB 7; Length 144;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
DB 36 CSATSNAC 43

RESULT 28
ADP31453
ID ADP31453 standard; protein; 144 AA.
XX
XX ADP31453;
AC
XX 12-AUG-2004 (first entry)
DT
XX
XX Human secreted protein SEQ ID #2220.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
OS
XX
XX WO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PF
XX
XX 29-AUG-2002; 2002US-0406576P.
PR
XX 29-AUG-2002; 2002US-0406579P.
PR
XX 29-AUG-2002; 2002US-0406585P.
PR
XX 29-AUG-2002; 2002US-0406588P.
PR
XX 29-AUG-2002; 2002US-0406608P.
PR
XX 29-AUG-2002; 2002US-0406611P.
PR
XX 29-AUG-2002; 2002US-0406612P.
PR
XX 29-AUG-2002; 2002US-0406616P.
PR
XX 29-AUG-2002; 2002US-0406640P.
PR
XX 29-AUG-2002; 2002US-0406642P.
PR
XX 29-AUG-2002; 2002US-0406646P.
PR
XX 29-AUG-2002; 2002US-0406653P.
PR
XX 29-AUG-2002; 2002US-0406655P.
PR
XX 29-AUG-2002; 2002US-0406666P.
PR
XX 17-SEP-2002; 2002US-0410946P.
PR
XX 17-SEP-2002; 2002US-0410947P.
PR
XX 17-SEP-2002; 2002US-0410948P.
PR
XX 17-SEP-2002; 2002US-0410949P.
PR
XX 17-SEP-2002; 2002US-0410953P.
PR
XX 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485224P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486860P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Halsehan L, Lannemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX MPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PI genetic, bacterial and viral diseases.
XX
XX
XX Claim 1; SEQ ID NO 3451; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
XX Sequence 144 AA;

Query Match 60.0%; Score 18; DB 8; Length 144;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
Db 20 CTAATTAC 27
RESULT 29
ADX90450 standard; protein; 144 AA.
XX
XX ADX90450;
XX
XX 21-APR-2005 (first entry)
XX
XX Plant full length insert polypeptide seqid 53114.
XX
XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
XX Unidentified.
XX
XX US2004034888-A1.
XX
XX 19-FEB-2004.
XX
XX 28-APR-2003; 2003US-00425114.
XX
XX 06-MAY-1999; 98US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
XX (LIU/J) LIU J.
PA (ZHOU/J) ZHOU Y.
PA (KOVA/) KOVALLIC D K.
PA (SCRE/) SCREEN S E.
PA (TABR/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
XX Liu J, Zhou Y, Kovallic DK, Screen SE, Tabaska JE, Cao Y;
PI
XX
XX MPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PI pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX
XX Claim 1; SEQ ID NO 53114; 15bp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or

CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.

XX Sequence 144 AA;

Query Match 60.0%; Score 18; DB 8; Length 144;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 4 CTSSAAC 11

RESULT 30

ID ABB45351 standard; protein; 150 AA.

XX ABB45351;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #30878.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Salmonella paratyphi.

PN WO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX WPI; 2003-029926/02.

DR N-PSDB; ACA49221.

PT New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 73275; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway;
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strings is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences

XX Sequence 150 AA;

Query Match 60.0%; Score 18; DB 6; Length 150;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 16 CATTASSC 23

RESULT 31

ID ABB41043 standard; protein; 153 AA.

XX ABB41043;

DT 08-FEB-2001 (first entry)

DE Human ORF807 polypeptide sequence SEQ ID NO:1614.

XX Human; open reading frame; ORF; detection; cytostatic; hepatotropic;

XX vulnary; antiproliferative; antiparkinsonian; neuroprotective;

XX anticonvulsant; osteopathic; antichratic; immunosuppressant; cardiac;

XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;

XX hypotensive; dermatological; immunosuppressive; antiinflammatory;

XX antiviral; antibacterial; antifungal; antineumatic; antihypertoid;

XX antianaemic; gene therapy; cancer; proliferative disorder; hypertension;

XX neurodegenerative disorder; osteoarthritis; graft vs host disease;

XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;

XX cholesterol ester storage; systemic lupus erythematosus; infection;

XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;

XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;

XX bone damage; cartilage damage; antiinflammatory disease; coagulation;

XX thrombosis; contraceptive.

XX Homo sapiens.

OS Homo sapiens.

XX WO200058473-A2.

PN 05-OCT-2000.

PF 31-MAR-2000; 2000WO-US008621.

PR 31-MAR-1999; 99US-0127607P.

PR 02-APR-1999; 99US-0127636P.

PR 05-APR-1999; 99US-0127728P.

PR 30-MAR-2000; 2000US-00540763.

XX (CURA-) CURAGEN CORP.

XX Shimketa RA, Leach M;

XX WPI; 2000-602362/57.

XX N-PSDB; AACT5252.

PT Novel nucleic acids and peptides derived from open reading frame X,

PT useful for treating e.g. cancers, proliferative disorders,

PT neurodegenerative disorders and cardiovascular disease.

XX Claim 11; Page 1302; 5507pp; English.
PS
XX AACT74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antiproliferative; antiparkinsonian; neurotropic; neuroprotective; osteopathic;
CC anticonvulsant; antirheumatic; immunosuppressant; immunostimulant;
CC cardiactant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antihemetic; antihypertensive; and antinausea. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ Sequence 153 AA;
Query Match 60.0%; Score 18; DB 3; Length 153;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
DB 78 CSSSASC 85
RESULT 32
ABP09251
ID ABP09251 standard; protein; 153 AA.
XX
XX ABP09251;
AC
XX
XX 25-JUN-2002 (first entry)
DT
XX
XX Human ORFX protein sequence SEQ ID NO:18484.
DE
XX
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
XX degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX hypertension; hypothyroidism; cholesterol ester storage disease;
XX immune deficiency; immune disorder; infectious disease;
XX autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
XX myasthenia gravis.
XX
XX Homo sapiens.
OS
XX
XX WO200192523-A2.
PN
XX
XX 06-DEC-2001.
PD
XX
XX 29-MAY-2001; 2001WO-US010836.
PF
XX
XX 30-MAY-2000; 2000US-0206132P.
PR
XX
XX 29-AUG-2000; 2000US-0228716P.
PA
XX
XX (CURA-) CURAGEN CORP.
PI
XX
XX Shimkets RA, Leach MD;
PI
XX
XX WPI; 2002-106308/14.
DR
XX
XX N-PSDB; ABR25003.
PT Novel human polypeptides and polynucleotides useful for diagnosing,

PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX
XX Disclosure; SEQ ID NO 18484; 1037pp; English.
XX
XX The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABR15762 to ABR2752 encode the human ORFX
CC proteins given in ABP00010 to ABP11500, ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
XX format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 153 AA;
Query Match 60.0%; Score 18; DB 5; Length 153;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
DB 78 CSSSASC 85
RESULT 33
ABO74421
ID ABO74421 standard; protein; 154 AA.
XX
XX ABO74421;
AC
XX
XX 29-JUL-2004 (first entry)
DT
XX
XX Pseudomonas aeruginosa polypeptide #6596.
DE
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
OS
XX
XX US6551795-B1.
PN
XX
XX 22-APR-2003.
PD
XX
XX 18-FEB-1999; 99US-00252991.
PF
XX
XX 18-FEB-1998; 98US-0074788P.
PR
XX
XX 27-JUL-1998; 98US-0094190P.
PA
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PI
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI
XX
XX WPI; 2003-615309/58.
DR
XX
XX N-PSDB; ABD07992.
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of

PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 23167; 455bp; English.
XX
XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-*P. aeruginosa* drugs, as templates for recombinant
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology. Sequences AB067826-
CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 154 AA;

Query Match 60.0%; Score 18; DB 7; Length 154;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
Db 120 CAAASAC 127

RESULT 34
AB077836
ID AB077836 standard; protein; 156 AA.
XX
AC AB077836;
XX
DT 29-JUL-2004 (first entry)
XX
DE *Pseudomonas aeruginosa* polypeptide #10011.
XX
XX Bacterial infection; *Pseudomonas aeruginosa* infection; antibacterial.
XX
OS *Pseudomonas aeruginosa*.
XX
PN US651795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
XX
DR N-PSDB; ABD11407.
XX
PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 26582; 455bp; English.
XX
CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,

CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-*P. aeruginosa* drugs, as templates for recombinant
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology. Sequences AB067826-
CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 156 AA;

Query Match 60.0%; Score 18; DB 7; Length 156;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
Db 40 CSASAMAC 47

RESULT 35
ADP30761
ID ADP30761 standard; protein; 159 AA.
XX
AC ADP30761;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1528.
XX
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; Inflammatory; Immune; human secreted protein.
XX
OS *Homo sapiens*.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX
PR 29-AUG-2002; 2002US-0406579P.
XX
PR 29-AUG-2002; 2002US-0406585P.
XX
PR 29-AUG-2002; 2002US-0406588P.
XX
PR 29-AUG-2002; 2002US-0406608P.
XX
PR 29-AUG-2002; 2002US-0406611P.
XX
PR 29-AUG-2002; 2002US-0406612P.
XX
PR 29-AUG-2002; 2002US-0406616P.
XX
PR 29-AUG-2002; 2002US-0406640P.
XX
PR 29-AUG-2002; 2002US-0406642P.
XX
PR 29-AUG-2002; 2002US-0406646P.
XX
PR 29-AUG-2002; 2002US-0406653P.
XX
PR 29-AUG-2002; 2002US-0406655P.
XX
PR 29-AUG-2002; 2002US-0406666P.
XX
PR 17-SEP-2002; 2002US-0410946P.
XX
PR 17-SEP-2002; 2002US-0410947P.
XX
PR 17-SEP-2002; 2002US-0410948P.
XX
PR 17-SEP-2002; 2002US-0410949P.
XX
PR 17-SEP-2002; 2002US-0410953P.
XX
PR 17-SEP-2002; 2002US-0410957P.
XX
PR 17-SEP-2002; 2002US-0410958P.
XX
PR 17-SEP-2002; 2002US-0410959P.
XX
PR 17-SEP-2002; 2002US-0410960P.
XX
PR 17-SEP-2002; 2002US-0410961P.
XX
PR 17-SEP-2002; 2002US-0410962P.
XX
PR 17-SEP-2002; 2002US-0411019P.
XX
PR 17-SEP-2002; 2002US-0411022P.
XX
PR 17-SEP-2002; 2002US-0411023P.
XX
PR 17-SEP-2002; 2002US-0411024P.

PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411010P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467206P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471306P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485252P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan U, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX
DR WPI; 2004-348438/32.
XX
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 2759; 428bp; English.
XX
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
XX available on WIPOMB and is not in the specification.
XX
SQ Sequence 159 AA;

Query Match 60.0%; Score 18; DB 8; Length 159;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 125 CAATATATC 132

RESULT 36
ADP30758
ID ADP30758 standard; protein; 162 AA.
XX
XX
AC ADP30758;
XX
XX
DT 12-AUG-2004 (first entry)
XX
XX
DE Human secreted protein SEQ ID #1525.
XX
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX
XX
OS Homo sapiens.
XX
XX
PN WO2004035732-A2.
XX
XX
PD 29-APR-2004.
XX
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411010P.
XX 17-SEP-2002; 2002US-0411111P.
XX 17-SEP-2002; 2002US-0463700P.
XX 18-APR-2003; 2003US-0463708P.
XX 18-APR-2003; 2003US-0463716P.
XX 18-APR-2003; 2003US-0463732P.
XX 02-MAY-2003; 2003US-0467199P.
XX 02-MAY-2003; 2003US-0467201P.
XX 02-MAY-2003; 2003US-0467203P.
XX 02-MAY-2003; 2003US-0467230P.
XX 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX PS Claim 1; SEQ ID NO 2756; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
XX SQ Sequence 162 AA;
XX
Query Match 60.0%; Score 18; DB 8; Length 162;
Best Local Similarity 25.0%; Pred. NO. 4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
DB 53 CAATTATC 60
RESULT 37
ADP30756
ID ADP30756 standard; protein; 162 AA.
XX
AC ADP30756;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1523.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX Cancer; Inflammatory; Immune; human secreted protein.
OS Homo sapiens.
XX
XX MO2004035732-AZ.
PN
XX
PD 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
PF
XX

PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411071P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467132P.
PR 02-MAY-2003; 2003US-0467139P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 02-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;

PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX	
DR	WPI; 2004-348438/32.
XX	
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases
XX	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT	genetic, bacterial and viral diseases.
XX	
PS	Claim 1; SEQ ID NO 2754; 428pp; English.
XX	
CC	The present invention relates to an isolated nucleic acid molecule
XX	encoding a polypeptide which is believed to be cytostatic,
CC	anti-inflammatory, immunosuppressive, antibacterial and virucidal. The
XX	composition and methods are useful for diagnosing, preventing and
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,
CC	immune, metabolic, genetic, bacterial and viral diseases. The present
CC	sequence represents a human secreted protein. The present sequence is
CC	available on WIPOMEB and is not in the specification.
XX	
SQ	Sequence 162 AA;
XX	
Query Match	60.0%; Score 18; DB 8; Length 162;
Best Local Similarity	25.0%; Pred. No. 4e+02; 6; Indels 0; Gaps 0;
Matches	2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Oy	4 CXXXXXXC 11
Db	53 CAATPAC 60
RESULT 38	
ID	ADP30757
ADP30757	ADP30757 standard; protein; 162 AA.
XX	
AC	ADP30757;
XX	
DT	12-AUG-2004 (first entry)
XX	
DE	Human secreted protein SEQ ID #1524.
XX	
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW	cancer; inflammatory; immune; human secreted protein.
XX	
OS	Homo sapiens.
XX	
PN	WO2004035732-A2.
XX	
PD	29-APR-2004.
XX	
PF	28-AUG-2003; 2003WO-US026780.
XX	
PR	29-AUG-2002; 2002US-0406576P.
XX	
PR	29-AUG-2002; 2002US-0406579P.
XX	
PR	29-AUG-2002; 2002US-0406585P.
XX	
PR	29-AUG-2002; 2002US-0406588P.
XX	
PR	29-AUG-2002; 2002US-0406611P.
XX	
PR	29-AUG-2002; 2002US-0406612P.
XX	
PR	29-AUG-2002; 2002US-0406616P.
XX	
PR	29-AUG-2002; 2002US-0406640P.
XX	
PR	29-AUG-2002; 2002US-0406642P.
XX	
PR	29-AUG-2002; 2002US-0406646P.
XX	
PR	29-AUG-2002; 2002US-0406653P.
XX	
PR	29-AUG-2002; 2002US-0406655P.
XX	
PR	29-AUG-2002; 2002US-0406666P.
XX	
PR	17-SEP-2002; 2002US-0410946P.
XX	
PR	17-SEP-2002; 2002US-0410947P.
XX	
PR	17-SEP-2002; 2002US-0410948P.
XX	
PR	17-SEP-2002; 2002US-0410953P.
XX	
PR	17-SEP-2002; 2002US-0410957P.
XX	
PR	17-SEP-2002; 2002US-0410958P.
XX	
PR	17-SEP-2002; 2002US-0410959P.

PR	17-SEP-2002;	2002US-04109610P.
PR	17-SEP-2002;	2002US-0410961P.
PR	17-SEP-2002;	2002US-0410962P.
PR	17-SEP-2002;	2002US-0411019P.
PR	17-SEP-2002;	2002US-0411022P.
PR	17-SEP-2002;	2002US-0411023P.
PR	17-SEP-2002;	2002US-0411024P.
PR	17-SEP-2002;	2002US-0411032P.
PR	17-SEP-2002;	2002US-0411035P.
PR	17-SEP-2002;	2002US-0411037P.
PR	17-SEP-2002;	2002US-0411041P.
PR	17-SEP-2002;	2002US-0411045P.
PR	17-SEP-2002;	2002US-0411046P.
PR	17-SEP-2002;	2002US-0411052P.
PR	17-SEP-2002;	2002US-0411055P.
PR	17-SEP-2002;	2002US-0411073P.
PR	17-SEP-2002;	2002US-0411082P.
PR	17-SEP-2002;	2002US-0411019P.
PR	17-SEP-2002;	2002US-0411111P.
PR	18-APR-2003;	2003US-0463700P.
PR	18-APR-2003;	2003US-0463708P.
PR	18-APR-2003;	2003US-0463716P.
PR	18-APR-2003;	2003US-0463732P.
PR	02-MAY-2003;	2003US-0467199P.
PR	02-MAY-2003;	2003US-0467201P.
PR	02-MAY-2003;	2003US-0467203P.
PR	19-MAY-2003;	2003US-0471306P.
PR	19-MAY-2003;	2003US-0471336P.
PR	22-MAY-2003;	2003US-0472420P.
PR	09-JUN-2003;	2003US-0472430P.
PR	09-JUN-2003;	2003US-0476609P.
PR	08-JUL-2003;	2003US-0485218P.
PR	08-JUL-2003;	2003US-0485223P.
PR	08-JUL-2003;	2003US-0485224P.
PR	08-JUL-2003;	2003US-0485325P.
PR	14-JUL-2003;	2003US-0486440P.
PR	14-JUL-2003;	2003US-0486480P.
PR	15-JUL-2003;	2003US-0486891P.
PR	15-JUL-2003;	2003US-0486960P.
PR	08-AUG-2003;	2003US-0493341P.
PR	08-AUG-2003;	2003US-0493370P.
PR	08-AUG-2003;	2003US-0493573P.
PR	08-AUG-2003;	2003US-0493577P.
PA	(FIVE) FIVE PRIME THERAPEUTICS INC.	
XX		
PI	Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;	
PI	Halsendack RF, Hsiang MM, Kothakota S, Haishan L, Linnemann T;	
PI	Pleence K, Wang Y, Wong UGP, Wu G, Zhang H;	
XX		
XX		
DR	WPI; 2004-348438/32.	
XX		
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	genetic, bacterial and viral diseases.	
XX		
PS	Claim 1; SEQ ID NO 2755; 428bp; English.	
XX		
CC	The present invention relates to an isolated nucleic acid molecule	
CC	encoding a polypeptide which is believed to be cytostatic,	
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	composition and methods are useful for diagnosing, preventing and	
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,	
CC	immune, metabolic, genetic, bacterial and viral diseases. The present	
CC	sequence represents a human secreted protein. The present sequence is	
CC	available on WIPOWEB and is not in the specification.	
XX		
SEQ	Sequence 162 AA;	

Best Local Similarity 25.0%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
Db 53 CATTATTC 60

RESULT 39

ADP30813
ID ADP30813 standard; protein; 162 AA.

XX ADP30813;
AC

XX 12-AUG-2004 (first entry)
DT

XX Human secreted protein SEQ ID #1580.
DE

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.
OS

XX WO2004035732-A2.
PN

XX 29-APR-2004.
PD

XX 28-AUG-2003; 2003WO-US026780.
PF

XX 29-AUG-2002; 2002US-0406576P.
PR

XX 29-AUG-2002; 2002US-0406579P.
PR

XX 29-AUG-2002; 2002US-0406585P.
PR

XX 29-AUG-2002; 2002US-0406588P.
PR

XX 29-AUG-2002; 2002US-0406608P.
PR

XX 29-AUG-2002; 2002US-0406611P.
PR

XX 29-AUG-2002; 2002US-0406612P.
PR

XX 29-AUG-2002; 2002US-0406616P.
PR

XX 29-AUG-2002; 2002US-0406642P.
PR

XX 29-AUG-2002; 2002US-0406646P.
PR

XX 29-AUG-2002; 2002US-0406653P.
PR

XX 29-AUG-2002; 2002US-0406655P.
PR

XX 29-AUG-2002; 2002US-0406666P.
PR

XX 17-SEP-2002; 2002US-0410944P.
PR

XX 17-SEP-2002; 2002US-0410947P.
PR

XX 17-SEP-2002; 2002US-0410948P.
PR

XX 17-SEP-2002; 2002US-0410949P.
PR

XX 17-SEP-2002; 2002US-0410953P.
PR

XX 17-SEP-2002; 2002US-0410957P.
PR

XX 17-SEP-2002; 2002US-0410959P.
PR

XX 17-SEP-2002; 2002US-0410960P.
PR

XX 17-SEP-2002; 2002US-0410961P.
PR

XX 17-SEP-2002; 2002US-0411019P.
PR

XX 17-SEP-2002; 2002US-0411022P.
PR

XX 17-SEP-2002; 2002US-0411023P.
PR

XX 17-SEP-2002; 2002US-0411024P.
PR

XX 17-SEP-2002; 2002US-0411032P.
PR

XX 17-SEP-2002; 2002US-0411035P.
PR

XX 17-SEP-2002; 2002US-0411037P.
PR

XX 17-SEP-2002; 2002US-0411041P.
PR

XX 17-SEP-2002; 2002US-0411045P.
PR

XX 17-SEP-2002; 2002US-0411046P.
PR

XX 17-SEP-2002; 2002US-0411048P.
PR

XX 17-SEP-2002; 2002US-0411052P.
PR

XX 17-SEP-2002; 2002US-0411055P.
PR

XX 17-SEP-2002; 2002US-0411073P.
PR

XX 17-SEP-2002; 2002US-0411082P.
PR

XX 17-SEP-2002; 2002US-0411101P.
PR

XX 17-SEP-2002; 2002US-0411111P.
PR

XX 18-APR-2003; 2003US-0463700P.
PR

XX 18-APR-2003; 2003US-0463708P.
PR

PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471366P.
PR 19-MAY-2003; 2003US-0471369P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PA Williams LT, Chu X, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Halenhan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 2811; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.

XX Sequence 162 AA;

XX Query Match

XX Best Local Similarity 25.0%; Score 18; DB 8; Length 162;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
Db 107 CATTTC 114

ADP31269

ID ADP31269 standard; protein; 165 AA.

XX ADP31269;
AC

XX 12-AUG-2004 (first entry)
DT

XX Human secreted protein SEQ ID #2036.
DE

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.
OS

XX	PN	MO2004035732-A2.	
XX	PD	29-APR-2004.	
XX	PF	28-AUG-2003; 2003MO-US0265780.	
XX	PR	29-AUG-2002; 2002US-0406576P.	
PR	29-AUG-2002; 2002US-0406579P.		
PR	29-AUG-2002; 2002US-0406585P.		
PR	29-AUG-2002; 2002US-0406588P.		
PR	29-AUG-2002; 2002US-0406608P.		
PR	29-AUG-2002; 2002US-0406611P.		
PR	29-AUG-2002; 2002US-0406612P.		
PR	29-AUG-2002; 2002US-0406616P.		
PR	29-AUG-2002; 2002US-0406640P.		
PR	29-AUG-2002; 2002US-0406642P.		
PR	29-AUG-2002; 2002US-0406646P.		
PR	29-AUG-2002; 2002US-0406653P.		
PR	29-AUG-2002; 2002US-0406655P.		
PR	29-AUG-2002; 2002US-0406662P.		
PR	29-AUG-2002; 2002US-0410957P.		
PR	17-SEP-2002; 2002US-0410958P.		
PR	17-SEP-2002; 2002US-0410959P.		
PR	17-SEP-2002; 2002US-0410960P.		
PR	17-SEP-2002; 2002US-0410961P.		
PR	17-SEP-2002; 2002US-0410962P.		
PR	17-SEP-2002; 2002US-0411019P.		
PR	17-SEP-2002; 2002US-0411022P.		
PR	17-SEP-2002; 2002US-0411023P.		
PR	17-SEP-2002; 2002US-0411024P.		
PR	17-SEP-2002; 2002US-0411032P.		
PR	17-SEP-2002; 2002US-0411035P.		
PR	17-SEP-2002; 2002US-0411037P.		
PR	17-SEP-2002; 2002US-0411041P.		
PR	17-SEP-2002; 2002US-0411045P.		
PR	17-SEP-2002; 2002US-0411046P.		
PR	17-SEP-2002; 2002US-0411048P.		
PR	17-SEP-2002; 2002US-0411052P.		
PR	17-SEP-2002; 2002US-0411055P.		
PR	17-SEP-2002; 2002US-0411073P.		
PR	17-SEP-2002; 2002US-0411082P.		
PR	17-SEP-2002; 2002US-0411111P.		
PR	18-APR-2003; 2003US-0463708P.		
PR	18-APR-2003; 2003US-0463708P.		
PR	18-APR-2003; 2003US-0463716P.		
PR	18-APR-2003; 2003US-0463732P.		
PR	02-MAY-2003; 2003US-0463719P.		
PR	02-MAY-2003; 2003US-0463720P.		
PR	02-MAY-2003; 2003US-0463720P.		
PR	02-MAY-2003; 2003US-0463720P.		
PR	19-MAY-2003; 2003US-0471336P.		
PR	19-MAY-2003; 2003US-0471336P.		
PR	22-MAY-2003; 2003US-0472420P.		
PR	22-MAY-2003; 2003US-0472420P.		
PR	09-JUN-2003; 2003US-0476640P.		
PR	09-JUN-2003; 2003US-0476641P.		
PR	08-JUL-2003; 2003US-0485218P.		
PR	08-JUL-2003; 2003US-0485223P.		
PR	08-JUL-2003; 2003US-0485224P.		
PR	08-JUL-2003; 2003US-0485335P.		
PR	14-JUL-2003; 2003US-0486446P.		
PR	14-JUL-2003; 2003US-0486480P.		
PR	15-JUL-2003; 2003US-0486891P.		
PR	15-JUL-2003; 2003US-0486960P.		
PR	08-AUG-2003; 2003US-0493341P.		
PR	08-AUG-2003; 2003US-0493347P.		

PR	08-AUG-2003; 2003US-0493573P.
PR	08-AUG-2003; 2003US-0493577P.
XX	(FIVE-) FIVE PRIME THERAPEUTICS INC.
XX	
P1	Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
P1	Halebeck RF, Huang WM, Kohakota S, Halehan L, Innemann T;
P1	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX	
DR	WPI; 2004-348436/32.
XX	
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT	genetic, bacterial and viral diseases.
XX	
PS	Claim 1; SEQ ID NO 3267; 428bp; English.
XX	
CC	The present invention relates to an isolated nucleic acid molecule
CC	encoding a polypeptide which is believed to be cytostatic,
CC	antitumorigenic, immunosuppressive, antibacterial and antiviral. The
CC	composition and methods are useful for diagnosing, preventing and
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,
CC	immune, metabolic, genetic, bacterial and viral diseases. The present
CC	sequence represents a human secreted protein. The present sequence is
CC	available on WIPWEB and is not in the specification.
SQ	
	Sequence 165 AA;
Query Match	60.0%; Score 18; DB 8; Length 165;
Best Local Similarity	25.0%; Pred. No. 4e+02; Mismatches 6; Indels 0; Gaps 0
Matches	2; Conservative 0; Mismatches 6; Indels 0; Gaps 0
OY	4 CXXXXXXC 11
Dd	62 CTATTTTC 69
RESULT 41	
ADP31099	
ID	ADP31099 standard; protein; 168 AA.
XX	
AC	ADP31099;
XX	
DT	12-AUG-2004 (first entry)
XX	
DE	Human secreted protein SEQ ID #1866.
XX	
KW	Cytostatic; Antitumorigenic; Immunosuppressive; Antibacterial; Virucide;
KW	cancer; inflammatory; immune; human secreted protein.
XX	
OS	Homo sapiens.
XX	
PN	WO2004035732-A2.
PD	
XX	29-APR-2004.
XX	
PF	28-AUG-2003; 2003WO-US026780.
XX	
XX	
PR	29-AUG-2002; 2002US-0406576P.
PR	29-AUG-2002; 2002US-0406579P.
PR	29-AUG-2002; 2002US-0406585P.
PR	29-AUG-2002; 2002US-0406588P.
PR	29-AUG-2002; 2002US-0406608P.
PR	29-AUG-2002; 2002US-0406611P.
PR	29-AUG-2002; 2002US-0406612P.
PR	29-AUG-2002; 2002US-0406616P.
PR	29-AUG-2002; 2002US-0406642P.
PR	29-AUG-2002; 2002US-0406643P.
PR	29-AUG-2002; 2002US-0406646P.
PR	29-AUG-2002; 2002US-0406653P.
PR	29-AUG-2002; 2002US-0406655P.
PR	29-AUG-2002; 2002US-0406666P.
PR	17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471305P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 14-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 15-JUL-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halebeck RF, Huang MM, Kotchoka S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3097; 428pp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC		immune, metabolic, genetic, bacterial and viral diseases. The present
CC		sequence represents a human secreted protein. The present sequence is
CC		available on WIPoWEB and is not in the specification.
XX		
SQ	Sequence 168 AA;	
OY	4 CXXXXXXC 11	
DB	25 CATTTTC 32	
RESULT 42		
ADP30794		
ID	ADP30794 standard; protein, 171 AA.	
XX		
AC	ADP30794;	
XX		
DT	12-AUG-2004 (first entry)	
XX		
DE	Human secreted protein SEQ ID #1561.	
XX		
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;	
XX	Cancer; inflammatory; immune; human secreted protein.	
OS	Homo sapiens.	
XX		
MN	WO2004035732-A2.	
XX		
PD	29-APR-2004.	
XX		
PF	28-AUG-2003; 2003WO-US026780.	
XX		
PR	29-AUG-2002; 2002US-0406576P.	
XX		
PR	29-AUG-2002; 2002US-0406579P.	
XX		
PR	29-AUG-2002; 2002US-0406585P.	
XX		
PR	29-AUG-2002; 2002US-0406588P.	
XX		
PR	29-AUG-2002; 2002US-0406608P.	
XX		
PR	29-AUG-2002; 2002US-0406611P.	
XX		
PR	29-AUG-2002; 2002US-0406612P.	
XX		
PR	29-AUG-2002; 2002US-0406616P.	
XX		
PR	29-AUG-2002; 2002US-0406642P.	
XX		
PR	29-AUG-2002; 2002US-0406646P.	
XX		
PR	29-AUG-2002; 2002US-0406653P.	
XX		
PR	29-AUG-2002; 2002US-0406655P.	
XX		
PR	29-AUG-2002; 2002US-0406666P.	
XX		
PR	17-SEP-2002; 2002US-0410946P.	
XX		
PR	17-SEP-2002; 2002US-0410947P.	
XX		
PR	17-SEP-2002; 2002US-0410948P.	
XX		
PR	17-SEP-2002; 2002US-0410949P.	
XX		
PR	17-SEP-2002; 2002US-0410949P.	
XX		
PR	17-SEP-2002; 2002US-0410953P.	
XX		
PR	17-SEP-2002; 2002US-0410957P.	
XX		
PR	17-SEP-2002; 2002US-0410958P.	
XX		
PR	17-SEP-2002; 2002US-0410959P.	
XX		
PR	17-SEP-2002; 2002US-0410960P.	
XX		
PR	17-SEP-2002; 2002US-0410961P.	
XX		
PR	17-SEP-2002; 2002US-0411019P.	
XX		
PR	17-SEP-2002; 2002US-0411037P.	
XX		
PR	17-SEP-2002; 2002US-0411037P.	
XX		
PR	17-SEP-2002; 2002US-0411041P.	
XX		
PR	17-SEP-2002; 2002US-0411045P.	
XX		
PR	17-SEP-2002; 2002US-0411046P.	
XX		
PR	17-SEP-2002; 2002US-0411048P.	
XX		
PR	17-SEP-2002; 2002US-0411052P.	

PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463722P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-047136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486466P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 2792; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 171 AA;
XX
Query Match 60.0%; Score 18; DB 8; Length 171;
Best Local Similarity 25.0%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
Db 161 CATTATTC 168
RESULT 43
ADP30793
ID ADP30793 standard; protein; 171 AA.
XX
AC ADP30793;
XX
DT 12-AUG-2004 (first entry)

XX
DE Human secreted protein SEQ ID #1560.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-047136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.

PR 08-JUL-2003; 2003US-048532SP.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486446P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D;
PI Halenbeck RE, Huang MM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2791; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
XX Sequence 171 AA;
SQ
Query Match 60.0%; Score 18; DB 8; Length 171;
Best Local Similarity 25.0%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
Db 161 CATTATTC 168
RESULT 44
AAG10193
ID AAG10193 standard; protein; 177 AA.
XX
XX AAG10193;
AC
XX
XX 17-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 8414.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141877P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147044P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147435P.
PR 09-AUG-1999; 99US-0148171P.
PR 10-AUG-1999; 99US-0148319P.
PR 11-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156586P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157177P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.

PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 60.0%; Score 18; DB 3; Length 177;
Best Local Similarity 25.0%; Pred. No. 46+02; Mismatches 0; Gaps 0;
Matches 2; Conservative 0; Indels 6;

QY 4 CXXXXXXC 11
Db 87 CXXXXSSC 94

RESULT 45

ID ADP30755 standard; protein; 177 AA.

XX ADP30755;

DT 12-AUG-2004 (first entry)

XX Human secreted protein SEQ ID #1522.

DE Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

XX WO2004035732-A2.

XX PD 29-APR-2004.

XX 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-047136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halembeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce Y, Wang Y, Mong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2753; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,

CC Immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
SQ Sequence 177 AA;
QY 4 CXXXXXXC 11
Db 145 CTATATTC 152
RESULT 46
ADP31682
ID ADP31682 standard; protein; 177 AA.
XX
XX ADP31682;
AC
XX
XX 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #2449.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 29-AUG-2002; 2002US-0406667P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-047136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485233P.
PR 08-JUL-2003; 2003US-0485244P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.

PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

PS Claim 1; SEQ ID NO 3680; 428pp; English.

XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.

XX Sequence 177 AA;

Query Match 60.0%; Score 18; DB 8; Length 177;
Best Local Similarity 25.0%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
Db 39 CAAAAATC 46

RESULT 47

ADTS6379

ID ADTS6379 standard; protein; 177 AA.

XX ADTS6379;

DT 13-JAN-2005 (first entry)

XX Plant polypeptide, SEQ ID 6456.

XX Plant; transgenic; cold tolerance; growth rate; drought tolerance;

XX disease resistance; galactomanan production; plant growth regulator;

XX heat tolerance; herbicide tolerance; lignin production;

XX extreme osmotic condition tolerance; pathogens resistance;

XX pest resistance; yield improvement; seed oil yield; seed protein yield.

XX Viridiplantae.

XX US2004216190-A1.

XX 28-OCT-2004.

XX 18-DEC-2003; 2003US-00739930.

XX 28-APR-2003; 2003US-00424599.

XX 28-APR-2003; 2003US-00425115.

XX (KOVA/) KOVALIC D K.

XX Kovalic DK;

XX WPI; 2004-757369/74.

PS Claim 2; SEQ ID NO 6456; 14pp; English.

XX The invention relates a recombinant DNA construct comprising a
XX polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomanan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant protein
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC segdata.uspto.gov/sequence.html?docID=20040216190.

XX Sequence 177 AA;

Query Match 60.0%; Score 18; DB 8; Length 177;

Best Local Similarity 25.0%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 87 CXXXXXXC 94

RESULT 48

ABM86593
ID ABM86593 standard; protein; 182 AA.

XX ABM86593;

DT 02-JUN-2005 (first entry)

DE Rice abiotic stress responsive polypeptide SEQ ID NO:4839.

KM abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.

OS Oryza sativa.

PN WO2003008540-A2.

XX 30-JAN-2003.

PD 21-JUN-2002; 2002WO-US019668.

XX 22-JUN-2001; 2001US-0300112P.

PR 24-AUG-2001; 2001US-0314662P.

PR 26-SEP-2001; 2001US-0325277P.

PR 21-NOV-2001; 2001US-0332132P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,

PI Moughamer T, Provart N, Rieke D, Zhu T;

XX MPI; 2003-248011/24.

PS Claim 1; SEQ ID NO 4839; 89pp; English.

XX The invention relates to novel abiotic stress responsive polynucleotides
CC and polypeptides. Also disclosed are vectors, expression cassettes, host
CC cells, and plants containing such polynucleotides. Also disclosed are
CC methods for using the polynucleotides and polypeptides to alter the
CC responsiveness of a plant to abiotic stress. The invention is useful in
CC agriculture. The nucleic acid is useful for determining whether a test
CC plant has been exposed to an abiotic stress condition. It is also useful
CC for selecting an agent that alters abiotic stress regulated
CC polynucleotide expression in a plant cell, and to identify a homolog or
CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC molecule and the polypeptide encoded by it are useful in altering the
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC stress, osmotic stress or any of their combinations. The present sequence
CC is used in the exemplification of the invention

XX Sequence 182 AA;

Query Match 60.0%; Score 18; DB 7; Length 182;

Best Local Similarity 25.0%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 15 CSTAAGAC 22

RESULT 49

ABM89559
ID ABM89559 standard; protein; 185 AA.

XX ABM89559;

DT 02-JUN-2005 (first entry)

DE Rice abiotic stress responsive polypeptide SEQ ID NO:7805.

KM abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.

OS Oryza sativa.

PN WO2003008540-A2.

XX 30-JAN-2003.

PD 21-JUN-2002; 2002WO-US019668.

XX 22-JUN-2001; 2001US-0300112P.

PR 24-AUG-2001; 2001US-0314662P.

PR 26-SEP-2001; 2001US-0325277P.

PR 21-NOV-2001; 2001US-0332132P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,

PI Moughamer T, Provart N, Rieke D, Zhu T;

XX MPI; 2003-248011/24.

PS Claim 1; SEQ ID NO 7805; 89pp; English.

XX The invention relates to novel abiotic stress responsive polynucleotides
CC and polypeptides. Also disclosed are vectors, expression cassettes, host
CC cells, and plants containing such polynucleotides. Also disclosed are
CC methods for using the polynucleotides and polypeptides to alter the
CC responsiveness of a plant to abiotic stress. The invention is useful in
CC agriculture. The nucleic acid is useful for determining whether a test
CC plant has been exposed to an abiotic stress condition. It is also useful
CC for selecting an agent that alters abiotic stress regulated
CC polynucleotide expression in a plant cell, and to identify a homolog or
CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC molecule and the polypeptide encoded by it are useful in altering the
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC stress, osmotic stress or any of their combinations. The present sequence
CC is used in the exemplification of the invention

XX Sequence 185 AA;

Query Match 60.0%; Score 18; DB 7; Length 185;

Best Local Similarity 25.0%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 148 CSTATTSC 155

RESULT 50

ADP31576
ID ADP31576 standard; protein; 192 AA.

XX ADP31576;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2343.

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 ; Search time 14.487 Seconds
(without alignments)
92.983 Million cell updates/sec

Title: US-09-932-322-2

Perfect score: 30
Sequence: 1 XXXXXXXXXXXXX 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 1000 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	60.0	177	2	F96719
2	18	60.0	355	2	A34597
3	18	60.0	367	2	C39590
4	18	60.0	492	2	S49147
5	18	60.0	843	1	JDVLJ1
6	18	60.0	843	1	JDVLJ2
7	18	60.0	1743	2	T26859
8	18	60.0	2543	2	T31687
9	17	56.7	26	2	A44036
10	17	56.7	89	2	S72742
11	17	56.7	101	2	T42262
12	17	56.7	120	2	S72979
13	17	56.7	122	2	TJ0150
14	17	56.7	123	2	S25574
15	17	56.7	135	1	RWUUV
16	17	56.7	139	2	A71123
17	17	56.7	145	2	H87436
18	17	56.7	182	2	T22305
19	17	56.7	203	2	E70641
20	17	56.7	212	2	H75277
21	17	56.7	224	2	A41128
22	17	56.7	230	2	S09778
23	17	56.7	237	2	F95411
24	17	56.7	238	1	JH0367
25	17	56.7	243	2	A35871
26	17	56.7	248	2	T47902
27	17	56.7	267	2	E87269
28	17	56.7	271	2	F70771
29	17	56.7	272	1	S72790

30	17	56.7	284	2	IS1172
31	17	56.7	292	2	C64795
32	17	56.7	292	2	D90710
33	17	56.7	292	2	H85560
34	17	56.7	297	2	A98294
35	17	56.7	297	2	AH2989
36	17	56.7	305	2	AE0857
37	17	56.7	305	2	G65056
38	17	56.7	307	2	B85925
39	17	56.7	307	2	A91080
40	17	56.7	309	2	T32376
41	17	56.7	347	2	B53522
42	17	56.7	348	2	T36890
43	17	56.7	350	2	C84548
44	17	56.7	352	2	B87507
45	17	56.7	365	2	G71287
46	17	56.7	376	2	D97175
47	17	56.7	400	2	C34443
48	17	56.7	400	2	AE1988
49	17	56.7	408	1	WMNVIA
50	17	56.7	408	2	A72869
51	17	56.7	412	2	AG3637
52	17	56.7	419	2	C40901
53	17	56.7	429	2	T16856
54	17	56.7	439	2	A36385
55	17	56.7	456	2	S20597
56	17	56.7	461	2	A31858
57	17	56.7	467	2	JC6150
58	17	56.7	471	2	A35480
59	17	56.7	471	2	S18446
60	17	56.7	489	2	S71955
61	17	56.7	492	1	WMAD55
62	17	56.7	539	2	H84640
63	17	56.7	541	2	T15299
64	17	56.7	544	2	S44814
65	17	56.7	547	1	SS1475
66	17	56.7	567	2	T49942
67	17	56.7	577	2	B37057
68	17	56.7	575	2	E86212
69	17	56.7	631	2	T00925
70	17	56.7	638	2	D86477
71	17	56.7	655	2	A59430
72	17	56.7	658	2	T19487
73	17	56.7	662	2	T17211
74	17	56.7	667	2	A48579
75	17	56.7	687	1	PVPMW
76	17	56.7	788	2	A37057
77	17	56.7	843	1	JDVLJ3
78	17	56.7	853	1	H70939
79	17	56.7	934	1	A34372
80	17	56.7	1028	2	T34360
81	17	56.7	1156	2	T23308
82	17	56.7	1279	2	A47363
83	17	56.7	1287	2	I46032
84	17	56.7	1345	2	S55669
85	17	56.7	1474	2	D88550
86	17	56.7	1522	2	H86380
87	17	56.7	1747	2	A45974
88	17	56.7	1857	2	S31212
89	17	56.7	1888	2	S78476
90	17	56.7	2395	1	S50820
91	17	56.7	2533	2	T28675
92	17	56.7	2533	2	T28674
93	17	56.7	2844	2	S28291
94	17	56.7	3124	2	A40020
95	17	56.7	3229	2	S27852
96	17	56.7	338	2	A47307
97	16	53.3	39	2	B41933
98	16	53.3	44	1	S38275
99	16	53.3	45	1	PFYZ
100	16	53.3	55	1	NTNMB2
101	16	53.3	55	2	E82522
102	16	53.3	57	2	A46654

transcription fact
c1g protein - Esc
2-(5'-triphosphor
2-(5'-triphosphor
hypothetical prote
permease [imported
conserved hypothet
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
20k cyclophilin -
probable iron-side
hypothetical prote
conserved hypothet
conserved hypothet
conserved hypothet
nifs family enzyme
nitrogenase cofact
nitrogenase cofact
immediate-early pr
early gene transac
cis-cis-muonate t
hypothetical prote
hypothetical prote
surface antigen se
ribonuclease inhib
ribonuclease-angio
Gram-negative bact
variant surface gl
variant surface gl
MDM2-like p53-bind
early B1B 55k prot
probable Rieske ir
hypothetical prote
F4B9.3 protein -
cytochrome P450 cy
hypothetical prote
integrin beta-6 ch
hypothetical prote
hypothetical prote
hypothetical prote
white protein - fr
integrin beta-6 ch
DNA-directed DNA p
probable n1rB prot
complement C6 prec
hypothetical prote
hypothetical prote
hypothetical prote
RNA helicase A - h
nuclear DNA helica
legument protein 7
protein ZC84.6 [im
protein T22F7.3 [i
collagen alpha 1(X
collagen alpha 1(X
collagen alpha 1(X
surface protein ty
alpha-51D immobill
alpha-51D-immobill
hypothetical prote
collagen alpha 1(X
probable cell-surf
heat-stable entero
mating pheromone E
toxin P1TX-II - sp
fulvicin C - Myxoc
neurotoxin B-II -
hypothetical prote
growth modulatory

103	16	53.3	57	2	B4654	176	16	53.3	187	2	S1618	zygote-specific pr
104	16	53.3	57	2	C4654	177	16	53.3	187	2	P90315	hypothetical prote
105	16	53.3	62	2	S54356	178	16	53.3	188	2	T3386	hypothetical prote
106	16	53.3	62	2	S54356	179	16	53.3	189	2	G70347	phenylecrylic acid
107	16	53.3	72	1	WMBP2P	180	16	53.3	190	2	G70347	kuntz type subtitl
108	16	53.3	73	2	D82519	181	16	53.3	190	2	A59416	kuntz type subtitl
109	16	53.3	79	2	G82811	182	16	53.3	190	2	A59418	kuntz type subtitl
110	16	53.3	83	2	S11843	183	16	53.3	195	2	E82857	plix protein XF003
111	16	53.3	83	2	AB3251	184	16	53.3	198	2	F83435	conserved hypothec
112	16	53.3	91	2	G90824	185	16	53.3	199	2	S35441	serine-cRNA ligase
113	16	53.3	92	2	D37057	186	16	53.3	201	2	S23324	gene achaeae prote
114	16	53.3	92	2	T30632	187	16	53.3	202	2	AC2875	hypothetical prote
115	16	53.3	100	2	A13337	188	16	53.3	202	2	T28939	hypothetical prote
116	16	53.3	104	2	B81045	189	16	53.3	205	2	T37947	hypothetical prote
117	16	53.3	109	2	S24253	190	16	53.3	207	2	E81418	hypothetical prote
118	16	53.3	114	2	H71018	191	16	53.3	208	2	S28675	hypothetical prote
119	16	53.3	115	2	B96664	192	16	53.3	208	2	P00278	hypothetical prote
120	16	53.3	118	1	PSNMK1	193	16	53.3	208	2	B64925	probable oxidoredu
121	16	53.3	118	1	PSNMK3	194	16	53.3	208	2	E90926	probable oxidoredu
122	16	53.3	118	2	D34860	195	16	53.3	208	2	A85775	urease accessory p
123	16	53.3	118	2	E34860	196	16	53.3	213	2	F75585	ribosomal protein
124	16	53.3	118	2	G34860	197	16	53.3	220	1	S71282	hypothetical prote
125	16	53.3	118	2	F34860	198	16	53.3	220	1	T21730	hypothetical prote
126	16	53.3	118	2	H34860	199	16	53.3	222	2	I39192	probable oxidoredu
127	16	53.3	119	2	PSNOA1	200	16	53.3	223	2	A65172	hypothetical prote
128	16	53.3	119	2	G97003	201	16	53.3	224	2	A36675	myogenin - mouse
129	16	53.3	120	2	E70424	202	16	53.3	225	2	B65127	type 4 preplin-11
130	16	53.3	125	2	JH0184	203	16	53.3	225	2	T49769	hypothetical prote
131	16	53.3	126	2	D72784	204	16	53.3	226	2	T33361	hypothetical prote
132	16	53.3	132	2	I39193	205	16	53.3	227	2	B35882	myogenin - chicken
133	16	53.3	133	2	B30242	206	16	53.3	229	1	TRBOTR	lyspsin (EC 3.4.21
134	16	53.3	135	2	D90847	207	16	53.3	229	2	T29203	hypothetical prote
135	16	53.3	136	2	I50554	208	16	53.3	231	1	TRBCTR	lyspsin (EC 3.4.21
136	16	53.3	138	2	E71883	209	16	53.3	231	2	T02765	glutathione transf
137	16	53.3	138	2	A90258	210	16	53.3	233	2	S71464	ribonuclease (EC 3
138	16	53.3	140	2	T49481	211	16	53.3	234	2	T52018	funarate reductase
139	16	53.3	140	2	T33359	212	16	53.3	237	2	S08073	cyclic nucleotide
140	16	53.3	141	2	T49380	213	16	53.3	237	2	T25152	hypothetical prote
141	16	53.3	141	2	PC1294	214	16	53.3	238	1	WZBR2	gene 2 protein - h
142	16	53.3	142	2	J50510	215	16	53.3	238	1	S31779	lyspsin (EC 3.4.21
143	16	53.3	143	1	QOEC32	216	16	53.3	240	2	S39047	lyspsin (EC 3.4.21
144	16	53.3	144	1	PSHUYF	217	16	53.3	240	2	H82289	conserved hypothec
145	16	53.3	144	2	JN0480	218	16	53.3	240	2	T25641	hypothetical prote
146	16	53.3	144	2	S01391	219	16	53.3	240	2	A43912	myogenin - Japanese
147	16	53.3	144	2	A60480	220	16	53.3	241	2	S39048	lyspsin (EC 3.4.21
148	16	53.3	145	1	PSKRF2U	221	16	53.3	242	2	C70895	hypothetical prote
149	16	53.3	145	1	S07983	222	16	53.3	242	2	T45795	hypothetical prote
150	16	53.3	145	2	S01390	223	16	53.3	245	2	T24565	hypothetical prote
151	16	53.3	145	2	S74292	224	16	53.3	247	2	S13813	lyspsin (EC 3.4.21
152	16	53.3	145	2	T42280	225	16	53.3	247	2	S05494	lyspsin (EC 3.4.21
153	16	53.3	145	2	T15608	226	16	53.3	248	2	S55067	lyspsin (EC 3.4.21
154	16	53.3	148	1	PSHU	227	16	53.3	248	2	T21786	hypothetical prote
155	16	53.3	149	2	C81959	228	16	53.3	248	2	F85067	hypothetical prote
156	16	53.3	150	2	C42744	229	16	53.3	249	2	T32060	hypothetical prote
157	16	53.3	151	2	T25047	230	16	53.3	254	2	G71545	probable deoxyoctu
158	16	53.3	153	2	S34255	231	16	53.3	254	2	A56447	CMR-2-keto-3-deoxy
159	16	53.3	157	2	S55956	232	16	53.3	259	2	T32542	hypothetical prote
160	16	53.3	160	2	JQ0542	233	16	53.3	259	2	T46881	hypothetical prote
161	16	53.3	166	2	F84449	234	16	53.3	261	2	F82728	succinate dehydrog
162	16	53.3	172	2	AB3514	235	16	53.3	262	2	G72003	probable outer mem
163	16	53.3	173	2	T14827	236	16	53.3	262	2	F86619	probable Omp (impo
164	16	53.3	174	2	AC0306	237	16	53.3	262	2	F87253	diethylactone hydro
165	16	53.3	177	2	G86917	238	16	53.3	262	2	A57391	TcSt-2 protein pre
166	16	53.3	177	2	D70654	239	16	53.3	265	2	A95401	protein (imported
167	16	53.3	178	2	G71195	240	16	53.3	266	2	S06462	beta-lactamase (EC
168	16	53.3	181	2	JN0342	241	16	53.3	266	2	S33163	beta-lactamase (EC
169	16	53.3	181	2	T15412	242	16	53.3	267	2	S22684	inhibitor of apopt
170	16	53.3	183	2	JX0311	243	16	53.3	268	2	T10304	apoptosis-inhibiti
171	16	53.3	183	2	I38923	244	16	53.3	268	2	A53989	insulin-like growt
172	16	53.3	183	2	A49313	245	16	53.3	272	1	A36082	hypothetical prote
173	16	53.3	183	2	S55599	246	16	53.3	274	2	T47933	chitinase (EC 3.2.
174	16	53.3	183	2	S55599	247	16	53.3	276	2	T09131	diaminopimelate ep
175	16	53.3	186	2	E82625	248	16	53.3	279	2	C70458	

249	16	53.3	280	2	T09504	LIM-protein FHL3,	322	16	53.3	366	2	A95382	probable iron upta
250	16	53.3	283	2	A64174	hypothetical prote	323	16	53.3	367	2	T49048	hypothetical prote
251	16	53.3	287	2	T15779	hypothetical prote	324	16	53.3	369	2	AE2345	phospho-N-acetylmu
252	16	53.3	288	2	T04401	endonuclease (EC 3	325	16	53.3	369	2	G81101	membrane-bound lye
253	16	53.3	288	2	T21790	hypothetical prote	326	16	53.3	370	2	T15213	hypothetical prote
254	16	53.3	288	2	T38075	conserved hypothet	327	16	53.3	370	2	B64632	hydrogenase expres
255	16	53.3	289	1	Q2AD2	early EIA 32K prot	328	16	53.3	370	2	F71883	hydrogenase expres
256	16	53.3	289	1	Q2AD5	early EIA 32K prot	329	16	53.3	371	2	A39234	opsin - blueottle
257	16	53.3	292	2	T23966	hypothetical prote	330	16	53.3	372	2	C97791	nife protein homol
258	16	53.3	293	2	T03883	hypothetical prote	331	16	53.3	373	1	OORF	rhodopsin - fruit
259	16	53.3	294	2	T23682	hypothetical prote	332	16	53.3	374	2	D83328	probable aminotran
260	16	53.3	296	2	S77427	2-hydroxy-6-oxohep	333	16	53.3	374	2	S40691	opsin rh1 - fruit
261	16	53.3	297	2	A10827	probable phosphog	334	16	53.3	374	2	T05660	hypothetical prote
262	16	53.3	297	2	T38161	hypothetical prote	335	16	53.3	374	2	T31842	zinc finger protei
263	16	53.3	298	2	AF0578	Citg protein (impo	336	16	53.3	376	2	B71652	iron-sulfur cofact
264	16	53.3	302	2	B71310	conserved hypothet	337	16	53.3	377	2	C89456	protein F55A.8 [i
265	16	53.3	303	2	T19289	hypothetical prote	338	16	53.3	378	2	A82406	cytochrome d ubiq
266	16	53.3	304	2	B83052	rRNA pseudouridine	339	16	53.3	379	2	S67856	Gumg protein - Xan
267	16	53.3	305	2	T49720	hypothetical prote	340	16	53.3	379	2	F71533	conserved hypothet
268	16	53.3	307	1	GSFP3	salivary glue prot	341	16	53.3	379	2	A81678	hypothetical prote
269	16	53.3	309	2	S41427	cysteine proteinas	342	16	53.3	381	2	A35300	G protein-coupled
270	16	53.3	309	2	S37028	exom protein - Rhi	343	16	53.3	383	2	I53870	Edg-1 orphan recep
271	16	53.3	309	2	S39957	succinoglycan bios	344	16	53.3	384	2	S68410	lipase Arab-1 - Ar
272	16	53.3	309	2	C95977	glucosyltransferas	345	16	53.3	384	2	F97711	hypothetical prote
273	16	53.3	311	2	S20085	transcription fact	346	16	53.3	384	2	T49084	hypothetical prote
274	16	53.3	312	1	QOBE07	US10 protein - hum	347	16	53.3	385	2	T04725	hypothetical prote
275	16	53.3	313	2	D85855	probable kinase ye	348	16	53.3	386	2	T06484	aspartate carboxy
276	16	53.3	313	2	B91011	probable kinase [i	349	16	53.3	387	2	C41880	nitrogenase cofact
277	16	53.3	313	2	B64985	hypothetical 33.6K	350	16	53.3	388	2	T16861	hypothetical prote
278	16	53.3	314	2	T27686	hypothetical prote	351	16	53.3	389	2	T23167	hypothetical prote
279	16	53.3	314	2	B43652	ferredoxin [2Fe-2S	352	16	53.3	391	1	U01626	attachment protein
280	16	53.3	318	2	B91177	probable hemin per	353	16	53.3	391	2	T08338	conserved hypothet
281	16	53.3	319	2	JC2252	chitinase (EC 3.2.	354	16	53.3	392	2	B47071	nitrogenase cofact
282	16	53.3	321	2	T02236	NBS-ILRR type resis	355	16	53.3	399	2	F82657	beta-ketoacyl - (ACP
283	16	53.3	322	2	S54806	chitinase (EC 3.2.	356	16	53.3	399	2	T21015	hypothetical prote
284	16	53.3	323	1	A48149	carcinoma-associat	357	16	53.3	404	2	A46274	HIV gp120-binding
285	16	53.3	323	2	T03614	chitinase (EC 3.2.	358	16	53.3	411	2	B97338	3-oxoacyl - (acyl -
286	16	53.3	323	2	S56777	probable membrane	359	16	53.3	412	2	S58105	Cu metalloregulato
287	16	53.3	328	2	C72370	cryptophan-tRNA [i	360	16	53.3	416	2	S27198	homeotic protein H
288	16	53.3	328	2	T36494	probable membrane	361	16	53.3	418	2	T15142	hypothetical prote
289	16	53.3	329	2	JC7935	ankyrin repeat and	362	16	53.3	424	2	UC5891	omega 6 desaturase
290	16	53.3	329	2	AB1327	diaminopimelate ep	363	16	53.3	425	2	B90113	seeryl-RNA synthet
291	16	53.3	329	2	AD1698	diaminopimelate ep	364	16	53.3	425	2	UC5909	AE33 protein - fru
292	16	53.3	330	1	UN0561	urokinase-type pla	365	16	53.3	425	2	A37912	thrombin receptor
293	16	53.3	330	2	I56100	complement factor	366	16	53.3	427	2	G71272	iron-sulfur cofact
294	16	53.3	330	2	C86023	hypothetical prote	367	16	53.3	429	2	T21113	hypothetical prote
295	16	53.3	330	2	T49644	hypothetical prote	368	16	53.3	431	2	T36682	probable secreted
296	16	53.3	330	2	T26004	homeotic protein E	369	16	53.3	432	2	C70381	large subunit of I
297	16	53.3	331	2	A30242	hypothetical prote	370	16	53.3	432	2	A43448	thrombin receptor
298	16	53.3	332	2	AD3535	glucosamine-transpor	371	16	53.3	435	2	T15143	hypothetical prote
299	16	53.3	334	2	H69214	probable thiamin A	372	16	53.3	439	2	T28841	hypothetical prote
300	16	53.3	335	2	B71361	homeobox protein A	373	16	53.3	442	2	T24196	hypothetical prote
301	16	53.3	335	2	G01448	effector cell prot	374	16	53.3	442	2	D84600	probable xylose
302	16	53.3	337	2	A53041	protein C50H11.5 [375	16	53.3	444	2	F86329	hypothetical prote
303	16	53.3	338	2	G88986	hypothetical prote	376	16	53.3	448	2	S41725	integrinase - Saccha
304	16	53.3	338	2	T06603	hypothetical prote	377	16	53.3	449	2	T26571	hypothetical prote
305	16	53.3	340	2	T48288	hypothetical prote	378	16	53.3	450	2	T17234	hypothetical prote
306	16	53.3	344	1	RMRC2	T-cell surface gly	379	16	53.3	450	2	AD3117	polygalacturonase
307	16	53.3	344	2	G82428	iron(III) ABC tran	380	16	53.3	453	2	A98170	hypothetical prote
308	16	53.3	344	2	S72723	dipeptide transpor	381	16	53.3	456	2	A31857	ribonuclease inhib
309	16	53.3	345	2	B85881	hypothetical prote	382	16	53.3	456	2	T37575	bacterial cell cyc
310	16	53.3	345	2	B65012	hypothetical prote	383	16	53.3	466	2	AD3477	malate dehydrogena
311	16	53.3	345	2	H91036	hypothetical prote	384	16	53.3	470	2	A40697	63K sperm flagella
312	16	53.3	345	2	A05279	surface antigen 51	385	16	53.3	474	2	S31712	beta-1,3-glucanase
313	16	53.3	347	2	S41638	T-cell surface gly	386	16	53.3	475	2	S85389	glutamate decarbox
314	16	53.3	348	2	JU0431	hypothetical 35.5K	387	16	53.3	475	2	T50686	peroxisomal Ca-dep
315	16	53.3	349	2	S53864	hypothetical prote	388	16	53.3	477	1	I38893	transcription fact
316	16	53.3	349	2	T16882	hypothetical prote	389	16	53.3	478	2	S31906	beta-1,3-glucanase
317	16	53.3	355	2	C96651	protein Ttp18.9 [i	390	16	53.3	481	2	T20309	hypothetical prote
318	16	53.3	358	2	T00576	probable GDSL-moti	391	16	53.3	486	2	JC7212	stretch-activated
319	16	53.3	359	2	D89007	protein F59B1.6 [i	392	16	53.3	486	2	T38087	hypothetical prote
320	16	53.3	359	2	T26813	hypothetical prote	393	16	53.3	487	2	F84727	hypothetical prote
321	16	53.3	364	2	B36313	hypothetical 42K p	394	16	53.3	488	1	JC2385	protein disulfide-

395	16	53.3	492	2	JC5169	alkaline nuclease	468	16	53.3	713	2	I50128	fibroblast growth
396	16	53.3	496	2	D83614	conserved hypotet	469	16	53.3	713	2	A35502	major surface-labe
397	16	53.3	498	2	AF0157	probable aldehyde	470	16	53.3	722	2	E71403	hypothetical prote
398	16	53.3	498	2	A48203	interleukin-14 pre	471	16	53.3	723	2	PN0509	intergrin beta-3 ch
399	16	53.3	499	2	B81914	probable periplasm	472	16	53.3	734	2	JC4861	ferriillin beta cha
400	16	53.3	504	1	A28807	protein disulfide-	473	16	53.3	735	2	G02937	ferriillin beta - cr
401	16	53.3	504	2	S41661	protein disulfide-	474	16	53.3	743	2	T02147	hypothetical prote
402	16	53.3	504	2	G02474	interferon regulat	475	16	53.3	746	2	S62365	SNR1-related prote
403	16	53.3	505	1	B69417	probable thymidine	476	16	53.3	752	2	T48574	hypothetical prote
404	16	53.3	505	1	S55507	protein disulfide-	477	16	53.3	752	2	T20871	hypothetical prote
405	16	53.3	505	2	S63994	protein disulfide-	478	16	53.3	753	2	B36268	platelet glycoprot
406	16	53.3	505	2	S68363	protein disulfide-	479	16	53.3	758	2	T48815	mixed-linked glucan
407	16	53.3	505	2	JC5704	protein disulfide-	480	16	53.3	761	2	A46193	88k B-26-specific
408	16	53.3	509	2	S46314	hydroxymethylgluta	481	16	53.3	763	2	E96693	probable terpene s
409	16	53.3	509	2	H70597	probable membrane	482	16	53.3	767	2	T07693	hypothetical prote
410	16	53.3	510	2	T48977	carboxypeptidase-1	483	16	53.3	769	2	S50966	probable membrane
411	16	53.3	510	2	H90787	Rtn-like protein [484	16	53.3	770	2	I40027	nitrite reductase
412	16	53.3	512	2	T37819	probable zinc meta	485	16	53.3	773	2	I46059	beta-1.integrin su
413	16	53.3	514	2	G01026	serine-tRNA ligase	486	16	53.3	778	2	A60798	platelet glycoprot
414	16	53.3	514	2	D56849	dopamine receptor-	487	16	53.3	781	2	S43534	integrin beta3 - C
415	16	53.3	519	2	T45764	hypothetical prote	488	16	53.3	781	2	A26547	platelet glycoprot
416	16	53.3	521	2	A99100	phosphatidylinosit	489	16	53.3	788	2	I77349	platelet glycoprot
417	16	53.3	521	2	T27192	hypothetical prote	490	16	53.3	788	2	I51530	integrin beta-3 su
418	16	53.3	522	2	S71821	probable interleuk	491	16	53.3	798	2	B27079	fibronectin recept
419	16	53.3	522	2	T13687	hypothetical prote	492	16	53.3	798	2	A28193	integrin beta-1 ch
420	16	53.3	528	2	T33599	hypothetical prote	493	16	53.3	798	2	B28193	integrin beta-1* C
421	16	53.3	533	2	H71492	probable hep-60 -	494	16	53.3	798	2	T22793	hypothetical prote
422	16	53.3	534	2	B81654	60 kDa chaperonin	495	16	53.3	799	2	A36308	integrin beta-5 ch
423	16	53.3	536	2	H85647	hypothetical prote	496	16	53.3	801	2	G82302	probable cellobios
424	16	53.3	547	1	A32244	60k cysteine-rich	497	16	53.3	805	2	I40029	nitrite reductase
425	16	53.3	547	2	S19607	alkaline phosphata	498	16	53.3	807	2	E83726	assimilatory nitri
426	16	53.3	547	2	B43584	60k cysteine-rich	499	16	53.3	810	2	B30848	plasma (EC 3.4.21
427	16	53.3	548	2	B84306	hypothetical prote	500	16	53.3	817	2	T25674	hypothetical prote
428	16	53.3	553	2	D71515	60k cysteine-rich	501	16	53.3	818	2	JC4058	fibroblast growth
429	16	53.3	554	2	C81671	60 kDa outer membr	502	16	53.3	822	2	S19947	fibroblast growth
430	16	53.3	554	2	T25288	60k cysteine-rich	503	16	53.3	822	2	B49151	fibroblast growth
431	16	53.3	556	1	S12602	60k cysteine-rich	504	16	53.3	822	2	T01622	probable salt-indu
432	16	53.3	556	2	A86560	60 kDa Cysteine-ri	505	16	53.3	823	2	D86165	protein F15K9.3 (f
433	16	53.3	557	1	B39439	60k cysteine-rich	506	16	53.3	829	2	JC4583	fibroblast growth
434	16	53.3	557	2	D84496	probable retroelem	507	16	53.3	834	2	T09369	hypothetical prote
435	16	53.3	557	2	A48434	variant-specific s	508	16	53.3	838	2	T04449	hypothetical prote
436	16	53.3	558	2	JC5204	60k cysteine-rich	509	16	53.3	849	2	E86306	Similar to tuftell
437	16	53.3	561	2	S63388	probable membrane	510	16	53.3	849	2	T01286	probable RNA-bind
438	16	53.3	569	2	D85354	hypothetical prote	511	16	53.3	863	1	S51789	VLDR receptor prec
439	16	53.3	573	2	T49610	related to rna bin	512	16	53.3	872	2	S33015	hypothetical prote
440	16	53.3	577	2	T28017	hypothetical prote	513	16	53.3	884	2	E75489	conserved hypotet
441	16	53.3	579	2	T24880	hypothetical prote	514	16	53.3	885	2	JN0339	N-methyl-D-asparta
442	16	53.3	589	2	T43310	fibulin-1D precurs	515	16	53.3	889	2	D71414	hypothetical prote
443	16	53.3	591	2	T39195	probable amino aci	516	16	53.3	897	2	S67283	hypothetical prote
444	16	53.3	600	2	B46642	DNA-directed DNA p	517	16	53.3	901	2	JN0337	N-methyl-D-asparta
445	16	53.3	601	2	B96744	unknown protein [1	518	16	53.3	905	2	T23229	hypothetical prote
446	16	53.3	604	2	T49577	hypothetical prote	519	16	53.3	906	2	JN0341	N-methyl-D-asparta
447	16	53.3	617	2	T49444	lustrin A related	520	16	53.3	907	2	A46296	N-methyl-D-asparta
448	16	53.3	631	2	S71508	dnak-type molecula	521	16	53.3	906	2	E96692	probable wall-asso
449	16	53.3	638	2	G02068	white homolog - hu	522	16	53.3	913	2	G64110	hypothetical prote
450	16	53.3	633	2	T25473	hypothetical prote	523	16	53.3	922	2	JN0340	N-methyl-D-asparta
451	16	53.3	655	2	T51792	hypothetical prote	524	16	53.3	922	2	JN0338	N-methyl-D-asparta
452	16	53.3	656	2	JC2005	integrin beta-5 ch	525	16	53.3	936	2	S64384	probable membrane
453	16	53.3	657	2	E72486	hypothetical prote	526	16	53.3	938	2	A46612	N-methyl-D-asparta
454	16	53.3	668	2	A42375	chemoreceptor prot	527	16	53.3	938	2	S21104	N-methyl-D-asparta
455	16	53.3	672	2	T12524	hypothetical prote	528	16	53.3	938	2	S19710	N-methyl-D-asparta
456	16	53.3	677	2	C42125	trophozoite cystel	529	16	53.3	943	2	A47551	N-methyl-D-asparta
457	16	53.3	680	2	PN0510	integrin beta-3 ch	530	16	53.3	946	2	T01460	hypothetical prote
458	16	53.3	686	2	S30075	feric reductase (531	16	53.3	959	2	JN0336	N-methyl-D-asparta
459	16	53.3	686	2	S43562	K08B5.3 protein -	532	16	53.3	964	2	S48404	probable membrane
460	16	53.3	689	2	T42760	fibulin, splice fo	533	16	53.3	991	1	PIBVCV	RNA I protein - cu
461	16	53.3	692	2	T32980	hypothetical prote	534	16	53.3	992	2	A31666	hypothetical prote
462	16	53.3	704	2	S21911	BRGcore-NS-23 prote	535	16	53.3	1007	2	T01437	hypothetical prote
463	16	53.3	704	2	T31227	cran protein homol	536	16	53.3	1025	2	B54718	dihydropyrimidine
464	16	53.3	706	2	A81848	probable TonB-depe	537	16	53.3	1025	2	A54718	probable membrane
465	16	53.3	707	2	G86894	hypothetical prote	538	16	53.3	1029	2	S56229	nitrite reductase
466	16	53.3	709	2	T29692	hypothetical prote	539	16	53.3	1044	2	T43155	nitrite reductase
467	16	53.3	712	2	T42990	fibulin 1, splice	540	16	53.3	1044	2	T43160	nitrite reductase

541	16	53.3	1046	2	A26838	prestalk protein p	614	15	50.0	27	2	PC4234	hypothetical prote
542	16	53.3	1069	2	S27922	nuclear antigen B8	615	15	50.0	31	2	B61014	defensin R-2 - rat
543	16	53.3	1084	2	T15616	hypothetical prote	616	15	50.0	34	2	B44336	neurotoxin Tx3-3 -
544	16	53.3	1089	2	T16283	hypothetical prote	617	15	50.0	35	2	S18224	filamentous hemagg
545	16	53.3	1101	2	T16840	hypothetical prote	618	15	50.0	40	2	UT0515	Ig heavy chain V-I
546	16	53.3	1129	2	A47511	protein-tyrosine k	619	15	50.0	42	2	A37907	thyrotropin beta c
547	16	53.3	1132	2	JW0091	Janus Kinase (BC 2	620	15	50.0	45	2	AC3539	hypothetical prote
548	16	53.3	1132	2	JC4127	protein-tyrosine k	621	15	50.0	46	2	B81072	hypothetical prote
549	16	53.3	1153	2	P84468	hypothetical prote	622	15	50.0	46	2	G53613	plectoxin IX - spi
550	16	53.3	1218	2	E84537	hypothetical prote	623	15	50.0	52	1	XBP1	bromelain inhibito
551	16	53.3	1235	2	C69165	hypothetical prote	624	15	50.0	52	2	S66609	bromelain inhibito
552	16	53.3	1236	2	T50904	Mg protoporphyrin	625	15	50.0	53	2	T10299	conotoxin-like pro
553	16	53.3	1274	2	T50904	transferrin-like p	626	15	50.0	53	2	T10405	conotoxin-like pro
554	16	53.3	1297	2	T52065	probable myb-relat	627	15	50.0	53	2	C72850	conotoxin homolog
555	16	53.3	1297	2	T30274	proteolysin - se	628	15	50.0	53	2	T30499	conotoxin-like pro
556	16	53.3	1307	2	G96711	unknown protein, 9	629	15	50.0	54	1	FEME	ferredoxin 214Fe-4
557	16	53.3	1316	2	G70535	probable rpoC proc	630	15	50.0	56	1	TIIPG	pancreatic secreto
558	16	53.3	1331	2	JB0352	mucin MUC5B, trach	631	15	50.0	57	1	NILLSA	nisin precursor -
559	16	53.3	1336	2	T23310	hypothetical prote	632	15	50.0	58	2	D82759	hypothetical prote
560	16	53.3	1347	2	T41321	BTB domain and Ank	633	15	50.0	59	2	T44147	B3 protein (import
561	16	53.3	1357	2	T16860	hypothetical prote	634	15	50.0	59	2	AE3319	hypothetical prote
562	16	53.3	1360	2	T33922	hypothetical prote	635	15	50.0	62	2	P82332	bacterioferritin-a
563	16	53.3	1373	2	JB0095	gastric mucin MUC5	636	15	50.0	63	2	S25772	testis-specific pr
564	16	53.3	1376	2	S63986	collagen alpha 5 c	637	15	50.0	63	2	A34484	metallothionein 1
565	16	53.3	1396	1	VCBB40	major capsid prote	638	15	50.0	64	2	AD2881	hypothetical prote
566	16	53.3	1397	2	T46354	hypothetical prote	639	15	50.0	66	2	T17934	hypothetical prote
567	16	53.3	1413	2	D84481	probable retroelem	640	15	50.0	68	2	B84730	hypothetical prote
568	16	53.3	1416	2	P88550	protein ZC84.1 [im	641	15	50.0	69	2	B97230	hypothetical prote
569	16	53.3	1444	2	T18856	angiogenesis inhib	642	15	50.0	70	2	S23316	hypothetical prote
570	16	53.3	1469	2	B36665	slit protein 2 pre	643	15	50.0	70	2	C83620	hypothetical prote
571	16	53.3	1475	2	A60026	cell communication	644	15	50.0	71	2	F69981	hypothetical prote
572	16	53.3	1509	2	T19486	hypothetical prote	645	15	50.0	71	2	T00008	copy number contro
573	16	53.3	1518	2	T28880	hypothetical prote	646	15	50.0	71	2	A59412	KGD-bearing platel
574	16	53.3	1523	2	T13953	MEG5 protein - ra	647	15	50.0	71	2	A59413	platelet-aggregati
575	16	53.3	1531	2	T42218	slit-1 protein hom	648	15	50.0	71	2	S13168	platelet-aggregati
576	16	53.3	1557	2	T28811	hypothetical prote	649	15	50.0	72	2	A43019	platelet aggregati
577	16	53.3	1557	2	T16210	hypothetical prote	650	15	50.0	72	2	A43019	platelet aggregati
578	16	53.3	1620	2	T27283	hypothetical prote	651	15	50.0	72	2	AB2528	platelet aggregati
579	16	53.3	1650	2	S53457	dominant autoantib	652	15	50.0	72	2	F43019	platelet aggregati
580	16	53.3	1700	2	S08167	Balbiani ring 3 pr	653	15	50.0	72	2	D43019	platelet aggregati
581	16	53.3	1772	2	A45532	major mezozoite bu	654	15	50.0	72	2	T43019	platelet aggregati
582	16	53.3	1784	2	T02844	cdcl6-related prot	655	15	50.0	72	2	B43019	platelet aggregati
583	16	53.3	1801	1	MMRPS	laminin beta-2 cha	656	15	50.0	72	2	B43020	platelet aggregati
584	16	53.3	1822	2	S63985	collagen alpha 2 c	657	15	50.0	73	2	T42302	hypothetical prote
585	16	53.3	1896	2	T01490	hypothetical prote	658	15	50.0	73	2	S78719	protein YER091c-a
586	16	53.3	2022	2	T48818	glucan 1,4-alpha-g	659	15	50.0	73	2	H43019	platelet aggregati
587	16	53.3	2049	2	T29227	hypothetical prote	660	15	50.0	73	2	C43019	platelet aggregati
588	16	53.3	2180	2	T29764	hypothetical prote	661	15	50.0	73	2	B43019	platelet aggregati
589	16	53.3	2182	2	T14320	calcineurin inhibi	662	15	50.0	73	2	A43020	platelet aggregati
590	16	53.3	2233	2	T28669	surface protein 51	663	15	50.0	73	2	A43019	platelet aggregati
591	16	53.3	2287	2	T23312	hypothetical prote	664	15	50.0	73	2	B40003	platelet aggregati
592	16	53.3	2399	2	H71879	toxoin-like outer m	665	15	50.0	73	2	A40003	platelet aggregati
593	16	53.3	2470	2	I50726	cation-independent	666	15	50.0	74	2	S33833	hypothetical prote
594	16	53.3	2529	2	B64635	toxoin-like outer m	667	15	50.0	74	2	H90820	hypothetical prote
595	16	53.3	2531	2	T16743	hypothetical prote	668	15	50.0	75	2	S19696	matrig phenomene B
596	16	53.3	2548	2	E59435	myosin IXA (import	669	15	50.0	76	1	TIAM3	protease inhibitor
597	16	53.3	2643	2	T29149	hypothetical prote	670	15	50.0	76	2	I83048	FSH beta-subunit -
598	16	53.3	2704	2	S09118	G surface protein	671	15	50.0	76	2	G69226	hypothetical prote
599	16	53.3	2718	2	A23475	G surface protein	672	15	50.0	76	2	A33034	neurogranin - bovi
600	16	53.3	2761	2	T21064	hypothetical prote	673	15	50.0	76	2	D44007	apoptoxin IV - tra
601	16	53.3	3133	2	S52093	hemocytin - silkw	674	15	50.0	76	2	C44007	apoptoxin VI - tra
602	16	53.3	3672	2	T23433	hypothetical prote	675	15	50.0	76	2	AH2120	hypothetical prote
603	16	53.3	3704	2	T37316	probable laminin a	676	15	50.0	77	2	JN0536	hypothetical 8.3K
604	16	53.3	4056	2	H96599	protein F14J16.10	677	15	50.0	78	2	D84620	hypothetical prote
605	16	53.3	4544	1	S02392	alpha-2-macroglobu	678	15	50.0	78	2	G98261	hypothetical prote
606	16	53.3	4545	1	S25111	alpha-2-macroglobu	679	15	50.0	79	2	C60767	retrovirus-related
607	16	53.3	4548	1	S00657	apoprotein(a) (BC	680	15	50.0	79	2	AH3369	hypothetical prote
608	16	53.3	4660	2	T42737	gp330 protein prec	681	15	50.0	80	2	I65235	testicular luteini
609	16	53.3	4753	1	A47437	LDL-receptor-relat	682	15	50.0	81	2	C65019	hypothetical prote
610	16	53.3	5107	2	T29144	partial CDS - Caen	683	15	50.0	81	2	AC1864	hypothetical prote
611	16	53.3	10797	2	T30192	probable peptide s	684	15	50.0	83	1	WBBP67	gene 18.7 protein
612	15	50.0	15	2	B59045	alpha-conotoxin Au	685	15	50.0	83	2	F35982	bifen alpha - puff
613	15	50.0	27	2	S55030	CAPS protein - ant	686	15	50.0	83	2	A34156	bitistatin - puff

687	15	50.0	85	1	W9BPB7	gene 19.2 protein	760	15	50.0	119	2	PH1554	Ig H chain V regio
688	15	50.0	85	2	A12765	hypothetical prote	761	15	50.0	119	2	A45574	13.3 kDa polypepti
689	15	50.0	86	2	S20471	class V zygote-spe	762	15	50.0	120	1	PSKFT3U	phospholipase A2 (
690	15	50.0	87	2	T18115	hypothetical prote	763	15	50.0	120	1	PSKFT3U	phospholipase A2 (
691	15	50.0	88	2	F89129	protein F52EL.6 (l	764	15	50.0	120	2	I50993	gonadotropin I bet
692	15	50.0	89	2	F90453	ferredoxin like pr	765	15	50.0	120	2	G71105	hypothetical prote
693	15	50.0	91	2	C38180	hypothetical prote	766	15	50.0	120	2	A12181	hypothetical prote
694	15	50.0	92	2	B90856	anther-specific pr	767	15	50.0	121	1	P5EGAC	phospholipase A2 (
695	15	50.0	94	2	T03285	prostatic steroid-	768	15	50.0	121	1	RWBEM3	surfact glycoprote
696	15	50.0	95	2	BOR3	3-mercaptopyruvate	769	15	50.0	121	1	PH1661	Ig heavy chain V r
697	15	50.0	96	2	A57483	alpha-(l)-3l-fucosy	770	15	50.0	122	1	PSABA	phospholipase A2 (
698	15	50.0	97	2	B97832	hypothetical prote	771	15	50.0	122	1	PSRSAA	phospholipase A2 (
699	15	50.0	98	2	T33302	hypothetical prote	772	15	50.0	122	1	PSRSAA	phospholipase A2 (
700	15	50.0	98	2	S42596	hypothetical prote	773	15	50.0	122	1	PSVIT	phospholipase A2 (
701	15	50.0	100	2	G87427	hypothetical prote	774	15	50.0	122	2	A53872	phospholipase A2 (
702	15	50.0	101	1	MMV211	11k protein - vacc	775	15	50.0	122	2	A25806	phospholipase A2 (
703	15	50.0	101	1	MMV212	11k protein - vacc	776	15	50.0	122	2	A44179	phospholipase A2 (
704	15	50.0	101	2	T30793	probable DNA-bindi	777	15	50.0	122	2	S46979	phospholipase A2 -
705	15	50.0	101	2	S73700	probable lipoprote	778	15	50.0	122	2	S62780	phospholipase A2 (
706	15	50.0	102	2	S50530	hypothetical prote	779	15	50.0	122	2	UX0063	phospholipase A2 (
707	15	50.0	102	2	S69884	hypothetical prote	780	15	50.0	122	2	S09314	phospholipase A2 (
708	15	50.0	102	2	C81889	hypothetical prote	781	15	50.0	122	2	T49039	response regulator
709	15	50.0	102	2	B82751	hypothetical prote	782	15	50.0	123	2	B53872	phospholipase A2 (
710	15	50.0	103	2	C69000	hypothetical prote	783	15	50.0	124	1	PSBGA2	phospholipase A2 (
711	15	50.0	103	2	T58936	hypothetical prote	784	15	50.0	124	1	NRCM	pancreatic ribonuc
712	15	50.0	104	2	D71035	hypothetical prote	785	15	50.0	124	1	NRCMB	pancreatic ribonuc
713	15	50.0	104	2	E81283	hypothetical prote	786	15	50.0	124	1	NRCMB	pancreatic ribonuc
714	15	50.0	104	2	S50911	metallochionein-2	787	15	50.0	124	1	A26535	phospholipase A2 (
715	15	50.0	106	2	S74084	foliitropin beta c	788	15	50.0	124	2	A59420	phospholipase A2 (
716	15	50.0	106	2	E69048	conserved hypotet	789	15	50.0	124	2	I51190	phospholipase A2 -
717	15	50.0	106	2	S32032	Sp17 protein precu	790	15	50.0	124	2	S25113	insulin-like growt
718	15	50.0	108	2	T45063	hypothetical prote	791	15	50.0	125	2	UX0052	phospholipase A2 (
719	15	50.0	109	2	PH1653	Ig heavy chain V r	792	15	50.0	125	2	S65624	phospholipase A(2)
720	15	50.0	109	2	T46171	hypothetical prote	793	15	50.0	125	2	AD0297	probable membrane
721	15	50.0	109	2	E70956	hypothetical prote	794	15	50.0	125	2	T28743	hypothetical prote
722	15	50.0	110	2	S55208	hypothetical prote	795	15	50.0	126	2	I36914	CD59 protein - bab
723	15	50.0	111	1	A29654	proteinase inhibit	796	15	50.0	126	2	F84240	hypothetical prote
724	15	50.0	111	1	S69911	Ig V-D-J region (R	797	15	50.0	126	2	G90350	conserved hypotet
725	15	50.0	111	2	H91187	hypothetical prote	798	15	50.0	127	2	S40034	killer toxin KP4 p
726	15	50.0	111	2	H85046	hypothetical prote	799	15	50.0	127	2	T10546	Rab11 protein homo
727	15	50.0	112	1	B29654	proteinase inhibit	800	15	50.0	128	1	RWHU59	surface glycoprote
728	15	50.0	113	1	S07092	gonadotropin beta	801	15	50.0	128	2	I36894	CD59 protein - gre
729	15	50.0	113	2	T36789	hypothetical prote	802	15	50.0	128	2	T35073	probable phosphori
730	15	50.0	113	2	S66924	probable membrane	803	15	50.0	128	2	AH3422	hypothetical membr
731	15	50.0	115	2	T45252	phosphatidoboyl-AMP	804	15	50.0	129	1	FTHUB	foliitropin beta c
732	15	50.0	117	2	A34860	phospholipase A2 (805	15	50.0	129	1	FTFGB	foliitropin beta c
733	15	50.0	117	2	B32416	phospholipase A2 (806	15	50.0	129	2	A23550	foliitropin beta c
734	15	50.0	117	2	A32416	phospholipase A2 (807	15	50.0	129	2	S73514	probable lipoprote
735	15	50.0	117	2	T26924	hypothetical prote	808	15	50.0	130	2	JC4526	foliitropin beta c
736	15	50.0	118	2	PSKFT1	phospholipase A2 (809	15	50.0	130	2	A33893	foliitropin beta c
737	15	50.0	118	1	PSKFT2	phospholipase A2 (810	15	50.0	131	2	I56894	complement regulat
738	15	50.0	118	1	PSKFT3	phospholipase A2 (811	15	50.0	131	2	A46298	pigment depositio
739	15	50.0	118	1	PSOXB	foliitropin beta c	812	15	50.0	131	2	A83350	hypothetical prote
740	15	50.0	118	1	FTHOB	phospholipase A2 (813	15	50.0	132	1	PSHOA	phospholipase A2 (
741	15	50.0	118	2	A32622	phospholipase A2 (814	15	50.0	132	2	I37143	agouti protein pre
742	15	50.0	118	2	B34860	phospholipase A2 (815	15	50.0	132	2	B83382	hypothetical prote
743	15	50.0	118	2	C34860	phospholipase A2 (816	15	50.0	133	1	PSOXG	phospholipase A2 (
744	15	50.0	118	2	S01801	luteolin beta chai	817	15	50.0	133	2	T1R2BR	crystin inhibitor
745	15	50.0	118	2	PN0141	luteolin beta chai	818	15	50.0	133	2	S14728	textilotoxin chain
746	15	50.0	118	2	PN0139	hypothetical prote	819	15	50.0	134	2	H84868	hypothetical prote
747	15	50.0	118	2	S61051	hypothetical prote	820	15	50.0	135	1	C64872	hypothetical prote
748	15	50.0	118	2	T49515	hypothetical prote	821	15	50.0	135	2	E83461	hypothetical prote
749	15	50.0	118	2	T32397	hypothetical prote	822	15	50.0	135	2	T15610	hypothetical prote
750	15	50.0	119	1	PSNJ2K	phospholipase A2 (823	15	50.0	136	2	AB0846	formate hydrogenly
751	15	50.0	119	1	PSNJ3K	phospholipase A2 (824	15	50.0	136	2	S08626	formate hydrogenly
752	15	50.0	119	1	PSN73K	phospholipase A2 (825	15	50.0	136	2	F91075	formate hydrogenly
753	15	50.0	119	1	PSNJ2W	phospholipase A2 (826	15	50.0	136	2	H82787	hypothetical prote
754	15	50.0	119	1	PSNJ3W	phospholipase A2 (827	15	50.0	136	2	T32396	hypothetical prote
755	15	50.0	119	1	PSOXA	phospholipase A2 (828	15	50.0	136	2	T36241	hypothetical prote
756	15	50.0	119	2	JN0403	phospholipase A2 (829	15	50.0	137	1	JC4877	hypothetical prote
757	15	50.0	119	2	S07528	phospholipase A2 (830	15	50.0	137	2	JC4710	beta-bungarotoxin
758	15	50.0	119	2	JX0115	pancreatic ribonuc	831	15	50.0	137	2	S37848	hypothetical prote
759	15	50.0	119	2	A61465	lutropin beta chai	832	15	50.0	137	2	T49243	hypothetical prote

833	15	50.0	137	2	T15609	hypothetical prote	906	15	50.0	146	2	T16342	phospholipase A2 (
834	15	50.0	137	2	A96914	uncharacterized pr	907	15	50.0	146	2	T01683	immunobillization ant
835	15	50.0	137	2	AE2763	transcription regu	908	15	50.0	147	1	PSKPA4	phospholipase A2 (
836	15	50.0	137	2	AF2790	conserved hypothet	909	15	50.0	147	2	T52489	ribonuclease 4 (EC
837	15	50.0	137	2	D97544	transcription regu	910	15	50.0	147	2	A53180	ribonuclease PL3 (
838	15	50.0	138	1	I51380	phospholipase A2 h	911	15	50.0	147	2	T50994	gonadotropin II be
839	15	50.0	138	1	PSRSAT	phospholipase A2 h	912	15	50.0	147	2	T20563	hypothetical prote
840	15	50.0	138	1	PSVTF	phospholipase A2	913	15	50.0	147	2	E82523	hypothetical prote
841	15	50.0	138	1	PSVIAA	phospholipase A2	914	15	50.0	147	2	F70216	hypothetical prote
842	15	50.0	138	1	PSVTAC	phospholipase A2	915	15	50.0	148	2	H82065	PTS system, nitrog
843	15	50.0	138	1	BGHU2	spermatid transit	916	15	50.0	148	2	T47180	Ig heavy chain var
844	15	50.0	138	2	A48168	phospholipase A2	917	15	50.0	149	2	T01676	cytochrome-c oxida
845	15	50.0	138	2	A46169	phospholipase A2	918	15	50.0	149	2	S70979	bifl. protein - Esc
846	15	50.0	138	2	F48188	phospholipase A2	919	15	50.0	149	2	AG2708	conserved hypothet
847	15	50.0	138	2	S29299	phospholipase A2	920	15	50.0	150	2	F69883	conserved hypothet
848	15	50.0	138	2	S59522	phospholipase A2	921	15	50.0	151	2	C90048	conserved hypothet
849	15	50.0	138	2	S10333	phospholipase A2	922	15	50.0	152	2	G81161	hypothetical prote
850	15	50.0	138	2	S15186	ammodictoxin B prec	923	15	50.0	153	2	T13054	hypothetical prote
851	15	50.0	138	2	S10992	ammodictoxin L prec	924	15	50.0	153	2	PN0103	hypothetical 17k p
852	15	50.0	138	2	JC1342	phospholipase A2	925	15	50.0	154	2	S35207	proteainase 7 - buf
853	15	50.0	138	2	I50098	phospholipase a2 -	926	15	50.0	154	2	S23320	hypothetical prote
854	15	50.0	138	2	A29479	thytotropin beta c	927	15	50.0	154	2	T27967	hypothetical prote
855	15	50.0	138	2	PC1211	envelope protein -	928	15	50.0	156	2	T31839	hypothetical prote
856	15	50.0	138	2	PC1212	envelope protein -	929	15	50.0	156	2	S74049	hypothetical prote
857	15	50.0	138	2	PC1205	envelope protein -	930	15	50.0	157	2	G65156	putative electron
858	15	50.0	138	2	PC1203	envelope protein -	931	15	50.0	157	2	S58017	hypothetical prote
859	15	50.0	138	2	PC1200	genome polyprotein	932	15	50.0	157	2	S58338	hypothetical prote
860	15	50.0	138	2	PC1210	envelope protein -	933	15	50.0	158	2	T52365	hypothetical prote
861	15	50.0	138	2	PC1207	envelope protein -	934	15	50.0	158	2	G86964	hypothetical prote
862	15	50.0	138	2	PC1197	genome polyprotein	935	15	50.0	158	2	T22829	hypothetical prote
863	15	50.0	138	2	PC1209	envelope protein -	936	15	50.0	159	2	S46822	hypothetical prote
864	15	50.0	138	2	PC1199	genome polyprotein	937	15	50.0	159	2	H81657	conserved hypothet
865	15	50.0	138	2	PC1201	envelope protein -	938	15	50.0	159	2	D71500	hypothetical prote
866	15	50.0	138	2	T14185	chitinase (EC 3.2.	939	15	50.0	159	2	T33966	hypothetical prote
867	15	50.0	138	2	T21792	hypothetical prote	940	15	50.0	160	2	T23995	hypothetical prote
868	15	50.0	139	1	PSIV	phospholipase A2	941	15	50.0	161	2	H64504	hypothetical prote
869	15	50.0	139	2	I52320	testicular luteini	942	15	50.0	161	2	S40711	hypothetical prote
870	15	50.0	139	2	A38612	insulin-like growt	943	15	50.0	161	2	T48146	hypothetical prote
871	15	50.0	139	2	S31612	beta-1,3-glucanase	944	15	50.0	162	2	H81058	disulfide bond for
872	15	50.0	139	2	AG0578	conserved hypothet	945	15	50.0	162	2	T50253	Vacuolar ATP synth
873	15	50.0	140	1	IIMSG1	interleukin-4 prec	946	15	50.0	162	2	T18188	hypothetical prote
874	15	50.0	140	1	HBFGC	hemoglobin beta ch	947	15	50.0	162	2	A72738	hypothetical prote
875	15	50.0	140	2	A48166	gonadotropin II be	948	15	50.0	162	2	E86387	18.7K hypothetical
876	15	50.0	140	2	H69405	hypothetical prote	949	15	50.0	163	2	A71469	hypothetical prote
877	15	50.0	140	2	T29914	hypothetical prote	950	15	50.0	164	1	A91016	ferredoxin-type pr
878	15	50.0	140	2	E87035	hypothetical prote	951	15	50.0	164	1	C85860	ferredoxin-type pr
879	15	50.0	141	1	B60826	gonadotropin beta	952	15	50.0	164	1	F64990	ferredoxin-type pr
880	15	50.0	141	1	UTBOB	lutropin beta chai	953	15	50.0	164	2	C35216	tumor suppressor p
881	15	50.0	141	1	UTBOB	lutropin beta chai	954	15	50.0	164	2	T84725	choriogonadotropin
882	15	50.0	141	1	UTRPA	lutropin beta chai	955	15	50.0	165	1	KTBAB	peptidoglycan asno
883	15	50.0	141	1	UTRPA	lutropin beta chai	956	15	50.0	165	1	KTBAB	hypothetical prote
884	15	50.0	141	1	UTSHB	lutropin beta chai	957	15	50.0	165	2	T40769	hypothetical prote
885	15	50.0	141	2	JC4527	lutinizng hormon	958	15	50.0	165	2	H84634	hypothetical prote
886	15	50.0	142	2	S54243	Ig mu heavy chain	959	15	50.0	165	2	E84346	placelact-derived g
887	15	50.0	142	2	AD2696	lipa protein (impo	960	15	50.0	165	2	S59899	chloroaurin chai
888	15	50.0	142	2	E96711	F245.7 (imported)	961	15	50.0	166	2	JN0248	hypothetical 18K p
889	15	50.0	143	2	T20951	hypothetical prote	962	15	50.0	166	2	JC1348	18.2K protein - ph
890	15	50.0	144	1	UTCAB	gonadotropin beta	963	15	50.0	166	2	A40783	probable plasmid s
891	15	50.0	144	1	JL0089	interleukin-9 prec	964	15	50.0	170	2	C96009	probable lipoprote
892	15	50.0	144	2	E88485	proteine F23F12.3	965	15	50.0	171	2	A10607	probable enzyme (l
893	15	50.0	144	2	T16824	hypothetical prote	966	15	50.0	171	2	D90747	Ybp1 protein precu
894	15	50.0	145	1	PSBOA	phospholipase A2	967	15	50.0	171	2	A64825	hypothetical prote
895	15	50.0	145	1	PSKRA2	phospholipase A2	968	15	50.0	171	2	T33072	hypothetical prote
896	15	50.0	145	2	T48093	phospholipase A2	969	15	50.0	171	2	E85046	hypothetical prote
897	15	50.0	145	2	I37231	beta-gonadotropin	970	15	50.0	172	2	T20210	hypothetical prote
898	15	50.0	146	1	PSDG	phospholipase A2	971	15	50.0	172	2	T10346	hypothetical prote
899	15	50.0	146	1	PSNJA'	phospholipase A2	972	15	50.0	173	1	B45932	atrial gland pepti
900	15	50.0	146	1	PSPCA	phospholipase A2	973	15	50.0	173	1	GOGAAA	atrial gland pepti
901	15	50.0	146	1	S34049	phospholipase A2	974	15	50.0	173	1	GOGABA	atrial gland pepti
902	15	50.0	146	1	S16763	gonadotropin beta	975	15	50.0	173	2	S62349	L71-3 protein - fr
903	15	50.0	146	2	A35493	phospholipase A2	976	15	50.0	173	2	H85597	probable enzyme yb
904	15	50.0	146	2	A33394	phospholipase A2	977	15	50.0	173	2	E97569	hypothetical prote
905	15	50.0	146	2	JU0283	phospholipase A2	978	15	50.0	174	2	T15176	hypothetical prote

979 15 50.0 174 2 A86358 similar to blue co
 980 15 50.0 175 2 T48541 hypothetical prote
 981 15 50.0 176 2 T18921 hypothetical prote
 982 15 50.0 177 1 AFMD8 phycoerythrin beta
 983 15 50.0 177 1 A55537 CDPdiacylglycerol-
 984 15 50.0 177 2 C71329 hypothetical prote
 985 15 50.0 178 2 T29345 hypothetical prote
 986 15 50.0 178 2 C84723 probable PHD-type
 987 15 50.0 178 2 D75390 NADH dehydrogenase
 988 15 50.0 180 2 S14019 hypothetical prote
 989 15 50.0 180 2 E84768 hypothetical prote
 990 15 50.0 180 2 F75607 arsenate reductase
 991 15 50.0 181 2 D86518 hypothetical prote
 992 15 50.0 181 2 C72104 hypothetical prote
 993 15 50.0 182 2 T45302 hypothetical prote
 994 15 50.0 182 2 T11906 NADH2 dehydrogenas
 995 15 50.0 183 2 A85561 hypothetical prote
 996 15 50.0 183 2 D64795 Apo-citrate lyase
 997 15 50.0 183 2 E90710 Apo-citrate lyase
 998 15 50.0 184 2 AG0736 hypothetical prote
 999 15 50.0 185 1 JC2394 phospholipase A2 1
 1000 15 50.0 185 2 C96808 protein F28K19.4 [

ALIGNMENTS

RESULT 1

unknown protein. 30275-30808 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C/Accession: F96719

R/Theologos, A.; Eckert, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Comp, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Hutzler, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Maritali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Xu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: F96719

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-177 <STO>

A/Cross-references: UNIPROT:P96C9L7; UNIPARC:UPI00000A1912; GB:AE05173; NID:9665540; PI

C/Genetics:

A/Gene: T6C23.4

A/Map position: 1

Query Match 60.0%; Score 18; DB 2; Length 177;
 Best Local Similarity 25.0%; Pred. No. 66;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
 DB 87 CSSSSSSC 94

RESULT 2

A34597 rhomboid protein - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster

C/Date: 06-Jul-1990 #sequence_revision 06-Jul-1990 #text_change 05-Oct-2004

C/Accession: A34597

R/Bier, E.; Jan, L.Y.; Jan, Y.N.

Genes Dev. 4, 190-203, 1990

A/Title: rhomboid, a gene required for dorsoventral axis establishment and peripheral ne

A/Reference number: A34597; MUID:90249726; PMID:2110920

A/Accession: A34597

A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-355 <BI>
 A/Cross-references: UNIPROT:P20350; UNIPARC:UPI000016BD27; GB:X52454; NID:910878; PIDN:C
 C/Genetics:
 A/Gene: FlyBase:ve
 A/Cross-references: FlyBase:FBgn0004635
 C/Superfamily: AARA protein
 C/Keywords: transmembrane protein

Query Match 60.0%; Score 18; DB 2; Length 355;
 Best Local Similarity 25.0%; Pred. No. 82;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
 DB 57 CSTASTC 64

RESULT 3

TPA-induced protein 11D - mouse

C/Species: Mus musculus (house mouse)

C/Date: 30-Aug-1991 #sequence_revision 30-Aug-1991 #text_change 09-Jul-2004

C/Accession: C39590

R/Varnum, B.C.; Ma, O.; Chi, T.; Fletcher, B.; Herschman, H.R.

Mol. Cell. Biol. 11, 1754-1758, 1991

A/Title: The Tis11 primary response gene is a member of a gene family that encodes prote

A/Reference number: A39590; MUID:91141531; PMID:1996120

A/Accession: C39590

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-367 <YAR>

A/Cross-references: UNIPROT:P23949; UNIPARC:UPI000002A130; GB:M58564; NID:9202066; PIDN

C/Keywords: DNA binding; zinc finger

Query Match 60.0%; Score 18; DB 2; Length 367;
 Best Local Similarity 25.0%; Pred. No. 82;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
 DB 273 CSSSASSC 280

RESULT 4

S49147 ERF-2 protein - human

C/Species: Homo sapiens (man)

C/Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004

C/Accession: S49147; I37372

R/Nie, X.F.; Maclean, R.N.; Kumar, V.; McKay, I.A.; Bustin, S.A.

submitted to the EMBL Data Library, April 1994

A/Reference number: S49147

A/Accession: S49147

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-492 <NIB>

A/Cross-references: UNIPROT:P47974; UNIPARC:UPI000003794E

R/Nie, X.F.; Maclean, R.N.; Kumar, V.; McKay, I.A.; Bustin, S.A.

Gene 152, 285-286, 1995

A/Title: ERF-2, the human homologue of the murine Tis11d early response gene.

A/Reference number: I37372; MUID:95137407; PMID:7835719

A/Accession: I37372

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-492 <RES>

A/Cross-references: UNIPARC:UPI000003794E; EMBL:X78992; NID:9509777; PIDN:CA55592.1; PI

C/Genetics:

A/Gene: GDB:BRF2; ERF2

A/Cross-references: GDB:511261

A/Map position: 13q12.3-13q12.3

Query Match 60.0%; Score 18; DB 2; Length 492;
Best Local Similarity 25.0%; Pred. No. 90;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 299 CXXXXXXC 306

RESULT 5

JDVLJ1
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adw, strain Japan/
C/Species: hepatitis B virus, HBV
C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C/Accession: D28925
R/Okamoto, H.; Tsuda, F.; Sakugawa, H.; Saastoeowigmo, R.I.; Imai, M.; Miyakawa, Y.; W
J. Gen. Virol. 69, 2573-2583, 1988
A/Title: Typing hepatitis B virus by homology in nucleotide sequence: comparison of surf
A/Reference number: J50253; MWID:89010694; PMID:3171552
A/Accession: D28925
A/Molecule type: DNA
A/Residues: 1-843 <OKA>
A/Cross-references: UNIPROT:P17394; UNIPARC:UPI00001297F9; GB:D00329; NID:G221497
C/Superfamily: hepatitis virus DNA-directed DNA polymerase
C/Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 60.0%; Score 18; DB 1; Length 843;
Best Local Similarity 25.0%; Pred. No. 1.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 268 CXXXXXXC 275

RESULT 6

JDVLJ2
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adw, strain Okinaw
C/Species: hepatitis B virus, HBV
C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C/Accession: E28925
R/Okamoto, H.; Tsuda, F.; Sakugawa, H.; Saastoeowigmo, R.I.; Imai, M.; Miyakawa, Y.; W
J. Gen. Virol. 69, 2573-2583, 1988
A/Title: Typing hepatitis B virus by homology in nucleotide sequence: comparison of surf
A/Reference number: J50253; MWID:89010694; PMID:3171552
A/Accession: E28925
A/Molecule type: DNA
A/Residues: 1-843 <OKA>
A/Cross-references: UNIPROT:P17395; UNIPARC:UPI00001297FC; GB:D00330; NID:G221498
C/Superfamily: hepatitis virus DNA-directed DNA polymerase
C/Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 60.0%; Score 18; DB 1; Length 843;
Best Local Similarity 25.0%; Pred. No. 1.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 268 CXXXXXXC 275

RESULT 7

T26859
hypothetical protein Y43F8B.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T26859
R/Ainscough, R.
submitted to the EMBL Data Library, October 1998
A/Reference number: Z20278
A/Accession: T26859
A/Status: preliminary; translated from GB/EMBL/DBD
A/Molecule type: DNA

A/Residues: 1-1743 <WIL>
A/Cross-references: UNIPROT:Q9XW5; UNIPARC:UPI0000078FD6; EMBL:AL032623; PIDN:CAA21511
A/Experimental source: clone Y43F8B
C/Genetics:
A/Gene: CESP:Y43F8B.3
A/Intons: 65/1; 92/2; 128/1; 229/1; 367/1; 422/1; 486/2; 523/1; 571/1; 628/1; 857/2; 9

Query Match 60.0%; Score 18; DB 2; Length 1743;
Best Local Similarity 25.0%; Pred. No. 1.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 1123 CXXXXXXC 1130

RESULT 8

T31687
surface antigen - Paramoecium primaurelia
C/Species: Paramoecium primaurelia
C/Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C/Accession: T31687
R/Bourgain-Guglielmetti, F.; Caron, F.
Journal of Eukaryot. Microbiol. 43, 303-314, 1996
A/Title: Molecular characterization of the D surface protein gene subfamily in Parameci
A/Reference number: Z21061; MWID:96313351; PMID:8768434
A/Accession: T31687
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-2543 <BOU>
A/Cross-references: UNIPROT:P90649; UNIPARC:UPI000007A903; EMBL:X96616; NID:G1235576; P
C/Genetics:
A/Superfamily: G surface protein

Query Match 60.0%; Score 18; DB 2; Length 2543;
Best Local Similarity 25.0%; Pred. No. 1.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 1961 CXXXXXXC 1968

RESULT 9

A44036
collagen alpha 1(XII) chain - bovine (fragment)
C/Species: Bos primigenius taurus (cattle)
C/Date: 27-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: A44036
R/Junstrum, G.P.; McDonough, A.M.; Marinkovich, M.P.; Keene, D.R.; Morris, N.P.; Burges
J. Biol. Chem. 267, 20087-20092, 1992
A/Title: Identification and partial purification of a large, variant form of type XII c
A/Reference number: A44036; MWID:93015874; PMID:1400326
A/Accession: A44036
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-26 <LUN>
A/Cross-references: UNIPROT:Q9TRM6; UNIPARC:UPI0000087EAB
A/Experimental source: skin
A/Note: sequence extracted from NCBI backbone (NCBI:P115795)

Query Match 56.7%; Score 17; DB 2; Length 26;
Best Local Similarity 25.0%; Pred. No. 89;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 10 CXXXXXXC 17

RESULT 10

S72742

B1177.F1.32 protein - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S72742
R:Smith, D.R.; Robison, K.
Submitted to the EMBL Data Library, November 1993
A:Description: Mycobacterium leprae cosmid B1177.
A:Reference number: S72694
A:Accession: S72742
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-89 <SMI>
A:Cross-references: UNIPROT:Q49652; UNIPARC:UPI0000086610; EMBL:U00011; NID:9466807; PID:G6netics:
A:Start codon: GTG

Query Match 56.7%; Score 17; DB 2; Length 89;
Best Local Similarity 25.0%; Pred. No. 1.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
| |
DB 71 CRTSTRAC 78

RESULT 11
T42262
hypothetical protein - phage SPPI
C:Species: phage SPPI
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T42262
R:Alonso, J.C.; Luder, G.; Stiege, A.C.; Chai, S.; Weise, F.; Trautner, T.A.
Gene 204, 201-212, 1997
A:Title: The complete nucleotide sequence and functional organization of Bacillus subtilis
A:Reference number: Z22137, MUID:98094274; PMID:9434185
A:Accession: T42262
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-101 <MO>
A:Cross-references: UNIPROT:Q48435; UNIPARC:UPI0000095544; EMBL:X97918; PID:CAA66572.1

Query Match 56.7%; Score 17; DB 2; Length 101;
Best Local Similarity 25.0%; Pred. No. 1.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
| |
DB 59 CCKTSSAC 66

RESULT 12
S72979
probable l-a-aminoadipyl-L-cystenyl-D-valine synthetase acvs - Mycobacterium leprae
N:Alternate names: B229 C2 179 protein
C:Species: Mycobacterium leprae
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S72979
R:Smith, D.R.; Robison, K.
Submitted to the EMBL Data Library, November 1993
A:Description: Mycobacterium leprae cosmid B229.
A:Reference number: S72588
A:Accession: S72979
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-120 <SMI>
A:Cross-references: UNIPROT:Q49859; UNIPARC:UPI00000852D7; EMBL:U00020; NID:9467102; PID:G6netics:
A:Gene: acvs
A:Start codon: GTG

Query Match 56.7%; Score 17; DB 2; Length 120;
Best Local Similarity 25.0%; Pred. No. 1.4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
| |
DB 4 CSYTSASC 11

RESULT 13
JQ0150
hypothetical 13k protein - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C:Accession: JQ0150
R:Kato, J.; Chu, L.; Kitanu, K.; Default, J.D.; Kimbara, K.; Chakrabarty, A.M.; Miera, T.
Gene 84, 31-38, 1989
A:Title: Nucleotide sequence of a regulatory region controlling alginate synthesis in Ps
A:Reference number: JQ0132; MUID:90108714; PMID:2514124
A:Accession: JQ0150
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-122 <KAT>
A:Cross-references: UNIPARC:UPI00001791D0
C:Superfamily: Pseudomonas aeruginosa hypothetical 13k protein

Query Match 56.7%; Score 17; DB 2; Length 122;
Best Local Similarity 25.0%; Pred. No. 1.4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
| |
DB 49 CGTATVAC 56

RESULT 14
S25574
IG heavy chain V region - human
C:Species: Homo sapiens (man)
C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 20-Jun-2000
C:Accession: S25574
R:Adelson, E.E.; Shackelford, P.G.; Quinn, A.; Carroll, W.L.
J. Immunol. 147, 1667-1674, 1991
A:Title: Restricted IG H chain V gene usage in the human antibody response to Haemophilu
A:Reference number: S25571; MUID:91349592; PMID:1908880
A:Accession: S25574
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-123 <ADD>
A:Cross-references: UNIPARC:UPI00001158BF; EMBL:X56527; NID:933569; PID:CAA39874.1; PID:
A:Note: the authors translated the codon GAA for residue 69 as Asp and TTA for residue 1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F;23-108/Domain: immunoglobulin homology <IMM>

Query Match 56.7%; Score 17; DB 2; Length 123;
Best Local Similarity 25.0%; Pred. No. 1.4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
| |
DB 113 CASSAYAC 120

RESULT 15
R02000
T-cell receptor beta chain precursor V region (Y735) - human
N:Alternate names: T-cell receptor beta-1 chain J-B1.2 segment
C:Species: Homo sapiens (man)
C:Date: 03-Aug-1984 #sequence_revision 03-Aug-1984 #text_change 09-Jul-2004
C:Accession: R02000; E24687
R:Yanagi, Y.; Yoshikai, Y.; Leggett, K.; Clark, S.P.; Alexander, I.; Mak, T.W.
Nature 308, 145-149, 1984
A:Title: A human T cell-specific cDNA clone encodes a protein having extensive homology
A:Reference number: A93324; MUID:84142269; PMID:6336315
A:Accession: A02000

A:Molecule type: mRNA
A:Residues: 1-135 <YAN>
A:Cross-references: UNIPROT:P01733; UNIPARC:UPI00001377D3; GB:K01571
A:Experimental source: clone Y735
R:Toyonaga, B.; Yoshikai, Y.; Vadasz, V.; Chin, B.; Mak, T.W.
Proc. Natl. Acad. Sci. U.S.A. 82, 8624-8628, 1985
A:Title: Organization and sequences of the diversity, joining, and constant region genes
A:Reference number: A94081; MUID:86094276; PMID:3866244
A:Accession: E24687
A:Molecule type: DNA
A:Residues: 121-135 <TOY>
A:Cross-references: UNIPARC:UPI0000113C76; GB:M4158; NID:g338844; PIDN:AAA60669.1; PID:
C:Genetics:
A:Gene: GDB:TCRB
A:Cross-references: GDB:120405; OMIM:186930
A:Map position: 7q35-7q35
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: glycoprotein; heterotrimer; receptor; T-cell
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-135/Product: T-cell receptor beta chain V region Y735 #status predicted <MAT>
F:35-113/Domain: immunoglobulin homology <IMM>
F:2-111/Disulfide bonds: #status predicted
F:90/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 56.7%; Score 17; DB 1; Length 135;
Best Local Similarity 25.0%; Pred. No. 1.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 111 CASSFSTC 118

RESULT 16
A71123
hypothetical protein PH0752 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C/Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C/Accession: A71123
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Onikubo, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: A71123
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-139 <KAW>
A:Cross-references: UNIPROT:O58508; UNIPARC:UPI0000062ECE; GB:AF000003; NID:g3236130; PI
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH0752

Query Match 56.7%; Score 17; DB 2; Length 139;
Best Local Similarity 25.0%; Pred. No. 1.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 81 CSSTVASC 88

RESULT 17
H87436
hypothetical protein CC1513 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: H87436
R:Netman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; Desoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolleva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87436
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-145 <TO>
A:Cross-references: UNIPROT:Q9A851; UNIPARC:UPI00000C73F0; GB:AE005673; NID:g13422892;
C:Genetics:
A:Gene: CC1513

Query Match 56.7%; Score 17; DB 2; Length 145;
Best Local Similarity 25.0%; Pred. No. 1.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 10 CAAABAC 17

RESULT 18
T22305
hypothetical protein F46C5.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T22305
R:Thomas, K.
submitted to the EMBL Data Library, September 1995
A:Reference number: Z19544
A:Accession: T22305
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-182 <WIL>
A:Cross-references: UNIPROT:P52881; UNIPARC:UPI000013A06E; EMBL:Z54281; PIDN:CAA91050.1
A:Experimental source: clone F46C5
C:Genetics:
A:Gene: CESP:F46C5.2
A:Map position: 2
A:Introns: 14/3; 50/1; 76/3; 145/2

Query Match 56.7%; Score 17; DB 2; Length 182;
Best Local Similarity 25.0%; Pred. No. 1.6e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 35 CTSGATTC 42

RESULT 19
E70641
hypothetical protein Rv0698 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: E70641
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Gantier, T.; Churcher, C.; Harris, D.; Gordon, S.
Comor, R.; Davies, R.; Pavlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: E70641
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-203 <COL>
A:Cross-references: UNIPROT:P95044; UNIPARC:UPI00000C14B5; GB:Z84395; GB:AL123456; NID:
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv0698

Query Match 56.7%; Score 17; DB 2; Length 203;
Best Local Similarity 25.0%; Pred. No. 1.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11
|
Db 51 CTTAHTC 58

RESULT 20
h75277
hypothetical protein - Deinococcus radiodurans (strain R1)
C/Species: Deinococcus radiodurans
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: H75277
R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A/Reference number: A75250; PMID:20036896; PMID:10567266
A/Accession: H75277
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-212 <WHI>
A/Cross-references: UNIPROT:Q9R88; UNIPARC:UPI00000CIA7; GB:AE002071; GB:AE000513; NID
C/Genetics:
A/Experimental source: strain R1
A/Map position: 1

Query Match 56.7%; Score 17; DB 2; Length 212;
Best Local Similarity 25.0%; Pred. No. 1.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11
|
Db 15 CTSASAC 22

RESULT 21
A41128
myogenin - human
N/Alternate names: myogenic determination factor Myf-4
C/Species: Homo sapiens (man)
C/Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 09-Jul-2004
C/Accession: A41128; S27295; S06948; S21978
R/Salminen, A.; Branun, T.; Buchberger, A.; Ueters, S.; Winter, B.; Arnold, H.H.
J. Cell Biol. 115, 905-917, 1991
A/Title: Transcription of the muscle regulatory gene MYF4 is regulated by serum component
A/Reference number: A41128; PMID:92064650; PMID:1659574
A/Accession: A41128
A/Molecule type: DNA
A/Residues: 1-224 <SAL>
A/Cross-references: UNIPROT:P15173; UNIPARC:UPI000012FBA4; EMBL:X62155
R/Branun, T.; Bober, E.; Buschhausen-Denker, G.; Koltz, S.; Grzeschik, K.H.; Arnold, H.H.
EMBO J. 9, 592, 1990
A/Reference number: S27295
A/Contents: extratum
A/Accession: S27295
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-224 <BRA>
A/Cross-references: UNIPARC:UPI000012FBA4
A/Note: this is a revision to the sequence from reference S06947
R/Branun, T.; Bober, E.; Buschhausen-Denker, G.; Koltz, S.; Grzeschik, K.H.; Arnold, H.H.
EMBO J. 8, 3617-3625, 1989
A/Title: Differential expression of myogenic determination genes in muscle cells: possib
A/Reference number: S06947; PMID:90059960; PMID:2583111
A/Accession: S06948
A/Molecule type: mRNA
A/Residues: 1-154, S06CPANALTAAPAAGVAGVHSSAPGTGGIICSLTQMPPTCTPSPSPSWTASQKMKCLMPQMKPC
A/Cross-references: UNIPARC:UPI000016AD95; EMBL:X17651; NID:934831; PIDN:CAA35641.1; PID
A/Note: this sequence has been revised in reference S27295
R/Arnold, H.H.
submitted to the EMBL Data Library, September 1991

A/Reference number: S21978
A/Accession: S21978
A/Molecule type: DNA
A/Residues: 1-223 <ARN>
A/Cross-references: UNIPARC:UPI000016AD96; EMBL:X62155; NID:934833; PIDN:CAA44080.1; PID
C/Genetics:
A/Genes: GDB:MYOG, MYF4
A/Cross-references: GDB:120210; OMIM:159980
A/Map position: 1q31-1q41
A/Introns: 157/3; 184/3
C/Superfamily: human myogenin
C/Keywords: DNA binding

Query Match 56.7%; Score 17; DB 2; Length 224;
Best Local Similarity 25.0%; Pred. No. 1.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11
|
Db 162 CSHSASC 169

RESULT 22
S09778
hypothetical protein UL16 precursor - human cytomegalovirus (strain AD169)
C/Species: human cytomegalovirus, human herpesvirus 5
A/Note: host Homo sapiens (man)
C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C/Accession: S09778
R/Chen, M.S.; Bankier, A.T.; Beck, S.; Bohm, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;
M.; Barrett, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A/Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A/Reference number: S09749; PMID:90269039; PMID:2161319
A/Accession: S09778
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-230 <CHE>
A/Cross-references: UNIPROT:P16757; UNIPARC:UPI0000137B56; EMBL:X17403; NID:959591; PIDN
A/Note: this sequence was submitted to the EMBL Data Library, December 1989
C/Superfamily: human cytomegalovirus hypothetical protein UL16
C/Keywords: glycoprotein; transmembrane protein
F/1-25/Domain: signal sequence #status predicted <SIG>
F/26-230/Product: hypothetical protein UL16 #status predicted <MAT>
F/189-206/Domain: transmembrane #status predicted <TM>
F/35,41,68,84,95,101,132,145/Binding site: carbohydrate (Aen) (covalent) #status predict

Query Match 56.7%; Score 17; DB 2; Length 230;
Best Local Similarity 25.0%; Pred. No. 1.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11
|
Db 18 CTSASAC 25

RESULT 23
P95411
hypothetical protein Sma221 [imported] - Sinorhizobium meliloti (strain 1021) magaplae
C/Species: Sinorhizobium meliloti
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C/Accession: P95411
R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barlow-Hubler, F.; Bow
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9889, 2001
A/Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo
A/Reference number: A95262; PMID:21396509; PMID:11481432
A/Accession: P95411
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-237 <MUR>
A/Cross-references: UNIPROT:Q92X56; UNIPARC:UPI00000CB344; GB:AE006469; PIDN:AAK65856.1
A/Experimental source: strain 1021, megaplasmid pSymA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, J.; Chain, D.; Chain, P.; Cowie, A.; Davies, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T. Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, J.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Sma2221
A:Genome: plasmid

Query Match 56.7%; Score 17; DB 2; Length 237;
Best Local Similarity 25.0%; Pred. No. 1.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
|
DB 76 CTAATTSC 83

RESULT 24

JH0367
ribonuclease (EC 3.1.27.-) Rh precursor - *Rhizopus niveus*
N:Alternate names: base nonspecific ribonuclease
C:Species: *Rhizopus niveus*
C>Date: 05-Mar-1993 #sequence_revision 31-Jan-1997 #text_change 31-Dec-2004
C:Accession: JH0367; A41449; B41449
R:Ohg1, K.; Horinuchi, H.; Watanabe, H.; Takagi, M.; Yano, K.; Irie, M.
J. Biochem. 109, 776-785, 1991
A:Title: Expression of RNase R from *Rhizopus niveus* in yeast and characterization of th
A:Reference number: JH0367; MUID:92011465; PMID:1655721
A:Accession: JH0367
A:Molecule type: DNA
A:Residues: 1-238 <OHG>
A:Cross-references: UNIPROT:P08056; UNIPARC:UPI00001344AC; DDBJ:D12476; DDBJ:D01125; NID
R:Horinuchi, H.; Yanai, K.; Takagi, M.; Yano, K.; Wakabayashi, E.; Sando, A.; Mine, S.; C
J. Biochem. 103, 408-418, 1988
A:Title: Primary structure of a base non-specific ribonuclease from *Rhizopus niveus*.
A:Reference number: A41449; MUID:88273061; PMID:3391995
A:Accession: A41449
A:Molecule type: DNA
A:Residues: 1-81, 'STR', 85-238 <HOR>
A:Cross-references: UNIPARC:UPI000016897A; GB:D00238; NID:G218044; PIDD:BAA00167.1; PID:
A:Note: the authors translated the codon TCA for residue 82 as Asn, CTC for residue 83 a
A:Accession: B41449
A:Molecule type: protein
A:Residues: 17-238 <HO2>
A:Cross-references: UNIPARC:UPI00001109AA
R:Kurikawa, H.; Nonaka, T.; Mitsui, Y.; Ohg1, K.; Irie, M.; Nakamura, K.T.
J. Mol. Biol. 259, 310-320, 1996
A:Title: The crystal structure of ribonuclease Rh from *Rhizopus niveus* at 2.0 Å resolution
A:Reference number: S62310; MUID:96150305; PMID:8531522
A:Contents: annotation; X-ray crystallography, 2.0 angstroms
C:Genetics:
A:introns: 21/1; 116/2; 151/3; 167/2
C:Function:
A:Description: hydrolyzes internal phosphodiester bonds of RNA to produce 3'-phosphomono
C:Superfamily: RNases
C:Keywords: hydrolase
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-238/Product: ribonuclease Rh #status experimental <MAT>
F:19-36,26-69,35-136,79-128,198-229/Disulfide bonds: #status experimental
F:52,121,125/Active site: His, Gln, His #status predicted
F:65,120,124/Binding site: substrate (TTP, His, Lys) #status predicted

Query Match 56.7%; Score 17; DB 1; Length 238;
Best Local Similarity 25.0%; Pred. No. 1.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
|

DB 19 CSSTALSC 26

RESULT 25

A35871
trypsin (EC 3.4.21.4) precursor, pancreatic - African clawed frog
C:Species: *Xenopus laevis* (African clawed frog)
C>Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 09-Jul-2004
C:Accession: A35871; S12117
R:Shi, Y.B.; Brown, D.D.
Genes Dev. 4, 1107-1113, 1990
A:Title: Developmental and thyroid hormone-dependent regulation of pancreatic genes in
A:Reference number: A35871; MUID:91007255; PMID:2210372
A:Accession: A35871
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-243 <SHI>
A:Cross-references: UNIPROT:P19799; UNIPARC:UPI00001376B0; EMBL:X53458; NID:965162; PID:
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; protein digestion; serine proteinase
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-20/Domain: activation peptide #status predicted <APR>
F:21-236/Domain: trypsin homology <TRY>
F:26-243/Product: trypsin I #status predicted <MAT>
F:27-157,45-61,129-230,136-203,168-182/Disulfide bonds: #status predicted
F:60,104,197/Active site: His, Asp, Ser #status predicted

Query Match 56.7%; Score 17; DB 2; Length 243;
Best Local Similarity 25.0%; Pred. No. 1.8e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
|
DB 129 CSAAGTSC 136

RESULT 26

T47902
hypothetical protein T4C21.320 - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: T47902
R:Choiane, N.; Robert, C.; Brotier, P.; Wincker, P.; Catolico, L.; Artiguenave, F.; S.
W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24479
A:Accession: T47902
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-248 <CHO>
A:Cross-references: UNIPROT:Q9LZX3; UNIPARC:UPI00009C7A5; EMBL:AL162295
A:Experimental source: cultivar Columbia; BAC clone T4C21
C:Genetics:
A:Map position: 3
A:introns: 61/3; 96/1; 119/1; 125/3; 143/2; 153/3; 181/1
A:Note: T4C21.320

Query Match 56.7%; Score 17; DB 2; Length 248;
Best Local Similarity 25.0%; Pred. No. 1.8e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
|
DB 191 CSSAKSC 198

RESULT 27

E87269
exopolysaccharide production protein Pse [imported] - *Caulobacter crescentus*
C:Species: *Caulobacter crescentus*
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 31-Dec-2004
C:Accession: E87269
R:Meriman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Kholmova, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A>Title: Complete Genome Sequence of *Caulobacter crescentus*.

A:Reference number: A87249; UID:21173698; PMID:11259647

A:Accession: E87269

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-267 <STO>

A:Cross-references: UNIPROT:Q9ABR0; UNIPARC:UPI000006CF59; GB:AE005673; NID:G13421281; F

C:Genetics:

A:Gene: CC0166

C:Superfamily: sugar transferase, ExoY type

Query Match 56.7%; Score 17; DB 2; Length 267;

Best local similarity 25.0%; Pred. No. 1.8e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

DB 20 CTTTSSDC 27

RESULT 28

F70771

probable glutamateracemase - *Mycobacterium tuberculosis* (strain H37RV)

C:Species: *Mycobacterium tuberculosis*

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: F70771

R:Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Raibaud, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skilton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A>Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome

A:Reference number: A70500; PMID:98295987; PMID:9634230

A:Accession: F70771

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-271 <COL>

A:Cross-references: UNIPROT:Q10626; UNIPARC:UPI000012PA0B; GB:Z73902; GB:AL123456; NID:G

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: mtrI

C:Superfamily: glutamate racemase

Query Match 56.7%; Score 17; DB 2; Length 271;

Best local similarity 25.0%; Pred. No. 1.8e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

DB 75 CNSASSAC 82

RESULT 29

S72790

glutamate racemase (EC 5.1.1.3) - *Mycobacterium leprae*

N:Alternate names: B1549 C2 210 protein

C:Species: *Mycobacterium leprae*

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: S72790

R:Smith, D.R.; Robison, K.

submitted to the EMBL Data Library, November 1993

A:Description: *Mycobacterium leprae* cosmid B1549.

A:Reference number: S72582

A:Accession: S72790

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-272 <SMI>

A:Cross-references: UNIPROT:P46705; UNIPARC:UPI000012PA0D; EMBL:U00014; NID:G466903; PID

C:Genetics:

A:Gene: mtrI

C:Superfamily: glutamate racemase

C:Keywords: Isomerase

Query Match 56.7%; Score 17; DB 1; Length 272;

Best local similarity 25.0%; Pred. No. 1.8e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

DB 75 CNTASAC 82

RESULT 30

I51172

transcription factor RCC/BBP-2 - bullfrog

C:Species: *Rana catesbeiana* (bullfrog)

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004

C:Accession: I51172

R:Chen, Y.; Hu, H.; Atkinson, B.G.

Dev. Genet. 15, 366-377, 1994

A>Title: Characterization and expression of C/EBP-like genes in the liver of *Rana catesbeiana*

A:Reference number: I51171; PMID:95008407; PMID:7923539

A:Accession: I51172

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-284 <CH>

A:Cross-references: UNIPROT:Q91295; UNIPARC:UPI000007BACC; EMBL:U08605; NID:G478890; PID

C:Superfamily: CCAT/enhancer-binding protein alpha

Query Match 56.7%; Score 17; DB 2; Length 284;

Best local similarity 25.0%; Pred. No. 1.8e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

DB 171 CSNTSSAC 178

RESULT 31

C64795

citG protein - *Escherichia coli* (strain K-12)

C:Species: *Escherichia coli*

C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004

C:Accession: C64795

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A>Title: The complete genome sequence of *Escherichia coli* K-12.

A:Reference number: A64720; PMID:97426617; PMID:9278503

A:Accession: C64795

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-292 <BLAT>

A:Cross-references: UNIPROT:P77231; UNIPARC:UPI0000127A40; GB:AE000166; GB:U00096; NID:G

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: citG

C:Superfamily: citG protein

Query Match 56.7%; Score 17; DB 2; Length 292;

Best local similarity 25.0%; Pred. No. 1.9e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

DB 142 CSTASFC 149

RESULT 32

D90710

2-(5'-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase (EC 4.2.-.-) *Escherichia coli*

C:Species: *Escherichia coli*

C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C:Accession: D90710
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
 Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A>Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
 A:Reference number: A9629; UID:2115631; PMID:11258796
 A:Accession: D90710
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-292 <HAY>
 A:Cross-references: UNIPROT:P58161; UNIPARC:UPI0000127A3F; GB:BA000007; PIDN:BA834075.1;
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: BCS0652
 C:Superfamily: citg protein
 C:Keywords: carbon-oxygen lyase

Query Match 56.7%; Score 17; DB 2; Length 292;
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
 DB 142 CSTAASFC 149

RESULT 33
 H85560
 2-(5''-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase (EC 4.2.-.-) citg - *Escherichia*
 C:Species: *Escherichia coli*
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C:Accession: H85560
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; DiMianta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: A85480; UID:21074935; PMID:11206551
 A:Accession: H85560
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-292 <STO>
 A:Cross-references: UNIPROT:P58161; UNIPARC:UPI0000127A3F; GB:AE005174; NID:G12513512; F
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: citg
 C:Superfamily: citg protein
 C:Keywords: carbon-oxygen lyase

Query Match 56.7%; Score 17; DB 2; Length 292;
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
 DB 142 CSTAASFC 149

RESULT 34
 A8294
 Hypothetical protein AGR_L_2619 [imported] - *Agrobacterium tumefaciens* (strain C58, Cere
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
 C:Accession: A98294
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; William, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2333-2328, 2001
 A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*
 A:Reference number: A97359; UID:21608551; PMID:11743194
 A:Accession: A98294
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-297 <KUR>
 A:Cross-references: UNIPROT:O8U53; UNIPARC:UPI00000D222C; GB:AE007870; PIDN:AAK89875.1;

C:Genetics:
 A:Gene: AGR_L_2619
 A:Map position: linear chromosome

Query Match 56.7%; Score 17; DB 2; Length 297;
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
 DB 151 CAAVAAC 158

RESULT 35
 AH2989
 perasease [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C:Accession: AH2989
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, J
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McClell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A>Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
 A:Reference number: AB2577; UID:21608550; PMID:11743193
 A:Accession: AH2989
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-297 <KUR>
 A:Cross-references: UNIPROT:O8U53; UNIPARC:UPI00000D222C; GB:AE008689; PIDN:AL44334.1,
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Adu3521
 A:Map position: linear chromosome

Query Match 56.7%; Score 17; DB 2; Length 297;
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
 DB 151 CAAVAAC 158

RESULT 36
 AE0857
 conserved hypothetical protein STY3065 [imported] - *Salmonella enterica* subsp. *enterica*
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh
 A>Note: this species has also been called *Salmonella typhi*
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AE0857
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 , S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.,
 A>Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
 A:Accession: AE0857
 A:Reference number: AB0502; UID:21534947; PMID:11677608
 A:Accession: AE0857
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-305 <PAR>
 A:Cross-references: UNIPARC:UPI000005A39C; GB:AL513382; PIDN:CAD06044.1; PID:G16504011,
 C:Genetics:
 A:Gene: STY3065
 C:Superfamily: *Escherichia coli* hypothetical protein b2755

Query Match 56.7%; Score 17; DB 2; Length 305;
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11
|
Db 179 CISAATSC 186

RESULT 37
G65056
hypothetical protein b2755 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: G65056
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: G65056
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-305 <BLAT>
A:Cross-references: UNIPROT:Q46936; UNIPARC:UPI000013B034; GB:AE000359; GB:U00096; NID:9
A:Experimental source: strain K-12, substrain MG1655
C:Superfamily: Escherichia coli hypothetical protein b2755

Query Match 56.7%; Score 17; DB 2; Length 305;
Best Local Similarity 25.0%; Pred. No. 1.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11
|
Db 179 CISAATSC 186

RESULT 38
B85925
hypothetical protein Z4064 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: B85925
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; DiMantano, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85925
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-307 <STO>
A:Cross-references: UNIPROT:Q8X7M3; UNIPARC:UPI000000D09E4; GB:AE005174; NID:912517212; F
A:Experimental source: strain O157:H7, substrain EDL933
C:Superfamily: Escherichia coli hypothetical protein b2755

Query Match 56.7%; Score 17; DB 2; Length 307;
Best Local Similarity 25.0%; Pred. No. 1.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11
|
Db 179 CISAATSC 186

RESULT 39
A91080
hypothetical protein ECG3609 [imported] - Escherichia coli (strain O157:H7, substrain R1
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: A91080
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gaerawa, N.; Yasunaga, T.; Kohara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gend

A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: A91080
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-307 <HAY>
A:Cross-references: UNIPROT:Q8X7M3; UNIPARC:UPI000000D09E4; GB:BA000007; PIDN:BAAB37032.1;
A:Experimental source: strain O157:H7, substrain R1MD 050952
C:Superfamily: Escherichia coli hypothetical protein b2755

Query Match 56.7%; Score 17; DB 2; Length 307;
Best Local Similarity 25.0%; Pred. No. 1.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11
|
Db 179 CISAATSC 186

RESULT 40
T32376
hypothetical protein K10F12.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T32376
R:Wohlmann, P.; Beck, C.
Submitted to the EMBL Data Library, September 1997
A:Description: The sequence of C. elegans cosmid K10F12.
A:Reference number: Z21157
A:Accession: T32376
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-309 <WOH>
A:Cross-references: UNIPROT:O17234; UNIPARC:UPI000007F612; EMBL:AF025462; PIDN:AAAB71002.
A:Experimental source: strain Bristol N2, clone K10F12
C:Genetics:
A:Gene: CSP:K10F12.4
A:Map position: 3
A:Introns: 31/3; 123/2; 196/3; 239/1

Query Match 56.7%; Score 17; DB 2; Length 309;
Best Local Similarity 25.0%; Pred. No. 1.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11
|
Db 7 CSTSALC 14

RESULT 41
B53522
20k cyclophilin - Toxoplasma gondii (fragment)
C:Species: Toxoplasma gondii
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: B53522
R:Higin, K.P.; Joiner, K.A.; Handschumacher, R.E.
J. Biol. Chem. 269, 9105-9112, 1994
A:Title: Isolation, cDNA sequences, and biochemical characterization of the major cyclo
A:Reference number: A53522; MUID:94179329; PMID:8132648
A:Accession: B53522
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-347 <HIG>
A:Cross-references: UNIPROT:Q26995; UNIPARC:UPI000007DB6A; GB:U04634; NID:9436958; PID:
F:178-346/Domain: cyclophilin homology <CYP>

Query Match 56.7%; Score 17; DB 2; Length 347;
Best Local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11
|
|

Db 48 CAGAAAC 55

RESULT 42

T36890
Probable iron-siderophore uptake system transmembrane component - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T36890
R:Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rejnders, M.A.
Submitted to the EMBL Data Library, August 1999
A/Reference number: Z21617
A/Accession: T36890
A/Status: preliminary; translated from GB/EMBL/DDBY
A/Molecule type: DNA
A/Residues: 1-348 <MUR>
A/Cross-references: UNIPROT:Q9S213; UNIPARC:UPI00000DB327; EMBL:AL109848; PIDN:CA852851.
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: SCOEDB:SC151.27c
C/Superfamily: vitamin B12 transport protein bluc

Query Match

Best Local Similarity 56.7%; Score 17; DB 2; Length 348;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
Db 262 CCAATPAC 269

RESULT 43

C84548
hypothetical protein At2g17120 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: C84548
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, F.; Vanaken, S.E.; Umayam, L.; Tallon, L.; Neus, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: C84548
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-350 <STO>
A/Cross-references: UNIPROT:Q23006; UNIPARC:UPI00000A22D; GB:AE002093; NID:92558660; PI
C/Genetics:
A/Gene: At2g17120
A/Map position: 2

Query Match

Best Local Similarity 56.7%; Score 17; DB 2; Length 350;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
Db 31 CCGSTSTC 38

RESULT 44

B87507
conserved hypothetical protein CC2083 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: B87507
R:Nierman, W.C.; Felldbylum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gilm, M.L.; Haft, D.H.; Kolon, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647

A/Accession: B87507
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-352 <STO>
A/Cross-references: UNIPROT:Q9A6U0; UNIPARC:UPI00000C75D3; GB:AE005673; NID:913423564;
C/Genetics:
A/Gene: CC2083

Query Match

Best Local Similarity 56.7%; Score 17; DB 2; Length 352;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
Db 11 CAALSAC 18

RESULT 45

G71287
conserved hypothetical protein TP0730 - syphilis spirochete
C/Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C/Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C/Accession: G71287
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwison, J.; Khailak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uetzerback, T.; McDye, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A/Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A/Reference number: A71250; MUID:98332770; PMID:9665876
A/Accession: G71287
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-365 <COL>
A/Cross-references: UNIPROT:Q83712; UNIPARC:UPI00000C0AA2; GB:AE001245; GB:AE000520; NID
A/Experimental source: strain Nichols
C/Genetics:
A/Gene: TP0730

Query Match

Best Local Similarity 56.7%; Score 17; DB 2; Length 365;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
Db 162 CAASSDC 169

RESULT 46

D97175
nifs family enzyme (cysteine desulfurase/cysteine sulfinatase desulfurase) [imported] - Clostridium acetobutylicum
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C/Accession: D97175
R:Molling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum ATCC824
A/Reference number: A96900; MUID:21359325; PMID:21359325
A/Accession: D97175
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-376 <KUR>
A/Cross-references: UNIPROT:Q97GY1; UNIPARC:UPI00000D7546; GB:AE001437; PIDN:AAK80191.1.
A/Experimental source: Clostridium acetobutylicum ATCC824
C/Genetics:
A/Gene: CAC2234
C/Superfamily: nitrogen fixation protein nifs

Query Match

Best Local Similarity 56.7%; Score 17; DB 2; Length 376;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

Db 319 CASAGSAC 326

RESULT 47

C34443
nitrogenase cofactor synthesis protein nifS - Anabaena sp.
N:Contains: L-cysteine sulfotransferase (EC 2.8.1.-)
C:Species: Anabaena sp.
C:Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 09-Jul-2004
C:Accession: C34443; B32361
R:Muligan, M.E.; Haselkorn, R.
J. Biol. Chem. 264, 19200-19207, 1989
A:Title: Nitrogen fixation (nif) genes of the cyanobacterium Anabaena species strain PCC
A:Reference number: A34443; PMID:90037054; PMID:2553733
A:Accession: C34443
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <MUL>
A:Cross-references: UNIPROT:P12623; UNIPARC:UPI000016E55C; GB:J05111; NID:g142034; PIDN:
R:Muligan, M.E.; Bulkema, W.J.; Haselkorn, R.
J. Bacteriol. 170, 4406-4410, 1988
A:Title: Bacterial-type ferredoxin genes in the nitrogen fixation regions of the cyanoba
A:Reference number: A91884; PMID:86314954; PMID:2842320
A:Accession: B32361
A:Molecule type: DNA
A:Residues: 1-7 <MU2>
A:Cross-references: UNIPARC:UPI0000178A6F
C:Superfamily: nitrogen fixation protein nifS
C:Keywords: phosphoprotein; pyridoxal phosphate; sulfotransferase
F:201/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted
F:324/Active site: Cys (cysteine persulfide intermediate) #status predicted

Query Match 56.7%; Score 17; DB 2; Length 400;
Best Local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 317 CASSGSAC 324

RESULT 48

AB1988
nitrogenase cofactor synthesis protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AB1988
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; PMID:21595285; PMID:11759840
A:Accession: AB1988
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <KUR>
A:Cross-references: UNIPROT:P12623; UNIPARC:UPI000013010C; GB:BA000019; PIDN:BAE73413.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: nifS
C:Superfamily: nitrogen fixation protein nifS

Query Match 56.7%; Score 17; DB 2; Length 400;
Best Local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 317 CASSGSAC 324

RESULT 49

MMNVIA
immediate-early protein IE-N - Autographa californica nuclear polyhedrosis virus
C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: A39150
R:Carson, D.D.; Summers, M.D.; Guarino, L.A.
Virology 182, 279-286, 1991
A:Title: Molecular analysis of a baculovirus regulatory gene.
A:Reference number: A39150; PMID:91220660; PMID:2024466
A:Accession: A39150
A:Molecule type: DNA
A:Residues: 1-408 <CAR>
A:Cross-references: UNIPROT:P24647; UNIPARC:UPI0000170DEB; GB:M59422; NID:g332437; PIDN:
C:Superfamily: AcMNPV immediate-early protein IE-N
C:Keywords: DNA binding; early protein; tandem repeat; transcription regulation
F:34-49/Region: 7-residue repeats
F:51-58/Region: 4-residue repeats
F:190-196/Region: glutamine-rich

Query Match 56.7%; Score 17; DB 1; Length 408;
Best Local Similarity 25.0%; Pred. No. 2.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 253 CSATSSDC 260

RESULT 50

A72869
early gene transactivator - Autographa californica nuclear polyhedrosis virus
C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV
A:Note: dsDNA virus
C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 09-Jul-2004
R:Byres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 586-605, 1994
A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.
A:Reference number: A72850; PMID:94303173; PMID:8030224
A:Accession: A72869
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-408 <AYR>
A:Cross-references: UNIPROT:P24647; UNIPARC:UPI00001387E9; GB:L22858; NID:g510708; PIDN:
C:Genetics:
A:Gene: Ac-IE-2
C:Superfamily: AcMNPV immediate-early protein IE-N

Query Match 56.7%; Score 17; DB 2; Length 408;
Best Local Similarity 25.0%; Pred. No. 2.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 253 CSATSSDC 260

Search completed: January 4, 2006, 16:09:57
Job time : 28.487 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 ; Search time 91.4261 Seconds
(without alignments)
108.037 Million cell updates/sec

Title: US-09-932-322-2

Perfect score: 30
Sequence: 1 XXXXXXXXXXXXX 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database :

1: uniprot_05.80:*
2: uniprot_sprot:*
3: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	60.0	69	2	Q9N669_GCAEN
2	18	60.0	69	2	Q9N669_GCAEN
3	18	60.0	69	2	Q9N669_GCAEN
4	18	60.0	69	2	Q9N669_GCAEN
5	18	60.0	69	2	Q9N669_GCAEN
6	18	60.0	69	2	Q9N669_GCAEN
7	18	60.0	69	2	Q9N669_GCAEN
8	18	60.0	69	2	Q9N669_GCAEN
9	18	60.0	69	2	Q9N669_GCAEN
10	18	60.0	69	2	Q9N669_GCAEN
11	18	60.0	69	2	Q9N669_GCAEN
12	18	60.0	69	2	Q9N669_GCAEN
13	18	60.0	69	2	Q9N669_GCAEN
14	18	60.0	69	2	Q9N669_GCAEN
15	18	60.0	69	2	Q9N669_GCAEN
16	18	60.0	69	2	Q9N669_GCAEN
17	18	60.0	69	2	Q9N669_GCAEN
18	18	60.0	69	2	Q9N669_GCAEN
19	18	60.0	69	2	Q9N669_GCAEN
20	18	60.0	69	2	Q9N669_GCAEN
21	18	60.0	69	2	Q9N669_GCAEN
22	18	60.0	69	2	Q9N669_GCAEN
23	18	60.0	69	2	Q9N669_GCAEN
24	18	60.0	69	2	Q9N669_GCAEN
25	18	60.0	69	2	Q9N669_GCAEN
26	18	60.0	69	2	Q9N669_GCAEN
27	18	60.0	69	2	Q9N669_GCAEN
28	18	60.0	69	2	Q9N669_GCAEN
29	18	60.0	69	2	Q9N669_GCAEN
30	18	60.0	69	2	Q9N669_GCAEN
31	18	60.0	69	2	Q9N669_GCAEN

32	18	60.0	445	2	Q8UUL4_ORYZA	Q8UUL4 oryzias lat
33	18	60.0	480	2	Q7SHC6_NEUTR	Q7shc6 neurospora
34	18	60.0	480	2	Q7SHY3_ASHGO	Q7shy3 ashya goss
35	18	60.0	492	1	TISD_HUMAN	P47974 homo sapien
36	18	60.0	494	2	Q53TB4_HUMAN	Q53tb4 homo sapien
37	18	60.0	526	2	Q6PD21_MOUSE	Q6pd21 mus muscucu
38	18	60.0	544	2	Q5B195_DROME	Q5b195 drosophila
39	18	60.0	545	2	Q91TR2_TUHY1	Q91tr2 tupaiid her
40	18	60.0	551	2	Q5Z8N9_ORYZA	Q5z8n9 oryza sativ
41	18	60.0	551	2	Q5Z8N9_ORYZA	Q5z8n9 oryza sativ
42	18	60.0	552	2	Q5AP87_CANAL	Q5ap87 candida alb
43	18	60.0	560	2	Q8U053_HPBVO	Q8u053 coxynebac
44	18	60.0	581	2	Q6N1S1_CORDI	Q6n1s1 cordybac
45	18	60.0	592	2	Q51VZ0_MACGR	Q51vz0 magnaporthe
46	18	60.0	606	2	Q6Z4E7_ORYZA	Q6z4e7 oryza sativ
47	18	60.0	691	2	Q4RVV9_TERNG	Q4rvv9 tetraodon n
48	18	60.0	705	2	Q7S136_NEUTR	Q7s136 neurospora
49	18	60.0	711	2	Q5RFE2_BRAKE	Q5rfe2 brachydanio
50	18	60.0	722	2	Q5AP84_CANAL	Q5ap84 candida alb
51	18	60.0	743	2	Q09517_HPBVO	Q09517 hepatitis b
52	18	60.0	769	2	Q8OGX5_HPBVO	Q8ogx5 hepatitis b
53	18	60.0	787	2	Q09511_HPBVO	Q09511 hepatitis b
54	18	60.0	795	2	Q8V1H4_HPBVO	Q8v1h4 hepatitis b
55	18	60.0	799	2	Q4T7H8_TERNG	Q4t7h8 tetraodon n
56	18	60.0	801	2	Q09504_HPBVO	Q09504 hepatitis b
57	18	60.0	801	2	Q09505_HPBVO	Q09505 hepatitis b
58	18	60.0	823	2	Q9QWN3_HPBVO	Q9qwn3 hepatitis b
59	18	60.0	836	2	Q8V1I3_HPBVO	Q8v1i3 hepatitis b
60	18	60.0	837	2	Q5O0U2_HPBVO	Q5o0u2 hepatitis b
61	18	60.0	843	1	DPOL_HPBVO	P17394 hepatitis b
62	18	60.0	843	1	DPOL_HPBVO	P17394 hepatitis b
63	18	60.0	843	1	DPOL_HPBVO	P17394 hepatitis b
64	18	60.0	843	2	Q09509_HPBVO	Q09509 hepatitis b
65	18	60.0	843	2	Q39877_HPBVO	Q39877 hepatitis b
66	18	60.0	843	2	Q39882_HPBVO	Q39882 hepatitis b
67	18	60.0	843	2	Q91527_HPBVO	Q91527 hepatitis b
68	18	60.0	843	2	Q5DW06_HPBVO	Q5dw06 hepatitis b
69	18	60.0	843	2	Q5DW09_HPBVO	Q5dw09 hepatitis b
70	18	60.0	843	2	Q5DW13_HPBVO	Q5dw13 hepatitis b
71	18	60.0	843	2	Q5O0T5_HPBVO	Q5o0t5 hepatitis b
72	18	60.0	843	2	Q67937_HPBVO	Q67937 hepatitis b
73	18	60.0	843	2	Q68RNI_HPBVO	Q68rni hepatitis b
74	18	60.0	843	2	Q68RP0_HPBVO	Q68rp0 hepatitis b
75	18	60.0	843	2	Q68RP3_HPBVO	Q68rp3 hepatitis b
76	18	60.0	843	2	Q68RO6_HPBVO	Q68ro6 hepatitis b
77	18	60.0	843	2	Q68RO9_HPBVO	Q68ro9 hepatitis b
78	18	60.0	843	2	Q68RR3_HPBVO	Q68rr3 hepatitis b
79	18	60.0	843	2	Q68RB4_HPBVO	Q68rb4 hepatitis b
80	18	60.0	843	2	Q765V6_HPBVO	Q765v6 hepatitis b
81	18	60.0	843	2	Q76B37_HPBVO	Q76b37 hepatitis b
82	18	60.0	843	2	Q77DP5_HPBVO	Q77dp5 hepatitis b
83	18	60.0	843	2	Q77DP8_HPBVO	Q77dp8 hepatitis b
84	18	60.0	843	2	Q77DP1_HPBVO	Q77dp1 hepatitis b
85	18	60.0	843	2	Q77DQ7_HPBVO	Q77dq7 hepatitis b
86	18	60.0	843	2	Q77DR0_HPBVO	Q77dr0 hepatitis b
87	18	60.0	843	2	Q77DR8_HPBVO	Q77dr8 hepatitis b
88	18	60.0	843	2	Q77DS1_HPBVO	Q77ds1 hepatitis b
89	18	60.0	843	2	Q77DT1_HPBVO	Q77dt1 hepatitis b
90	18	60.0	843	2	Q77HP5_HPBVO	Q77hp5 hepatitis b
91	18	60.0	843	2	Q77HQ0_HPBVO	Q77hq0 hepatitis b
92	18	60.0	843	2	Q77HT0_HPBVO	Q77ht0 hepatitis b
93	18	60.0	843	2	Q805G6_HPBVO	Q805g6 hepatitis b
94	18	60.0	843	2	Q80GV8_HPBVO	Q80gv8 hepatitis b
95	18	60.0	843	2	Q80H00_HPBVO	Q80h00 hepatitis b
96	18	60.0	843	2	Q80H20_HPBVO	Q80h20 hepatitis b
97	18	60.0	843	2	Q80H48_HPBVO	Q80h48 hepatitis b
98	18	60.0	843	2	Q80J60_HPBVO	Q80j60 hepatitis b
99	18	60.0	843	2	Q80J72_HPBVO	Q80j72 hepatitis b
100	18	60.0	843	2	Q80J80_HPBVO	Q80j80 hepatitis b
101	18	60.0	843	2	Q80J83_HPBVO	Q80j83 hepatitis b
102	18	60.0	843	2	Q80M08_HPBVO	Q80m08 hepatitis b
103	18	60.0	843	2	Q81116_HPBVO	Q81116 hepatitis b
104	18	60.0	843	2	Q81120_HPBVO	Q81120 hepatitis b


```
981 16 53.3 122 2 063fj9_BACCC 063fj9 bacillus ce
982 16 53.3 122 2 05UVZ1_GADBN 05UVZ1 human adeno
983 16 53.3 122 2 071923_GADB5 071923 human adeno
984 16 53.3 122 2 08B8S0_GADBN 08B8S0 human adeno
985 16 53.3 123 1 IL21_PERMA 080X92 peromyscus
986 16 53.3 123 2 07PPJ8_ANOGA 07PPJ8 anopheles g
987 16 53.3 123 2 08JG94_GINCI 08JG94 ginglymasto
988 16 53.3 124 2 08PVQ6_METMA 08PVQ6 methanobarc
989 16 53.3 124 2 07R266_GIALA 07R266 giardia lam
990 16 53.3 124 2 016121_TENMO 016121 tenebrio mo
991 16 53.3 124 2 062710_CAEBR 062710 caenorhabdi
992 16 53.3 124 2 09U744_TENMO 09U744 tenebrio mo
993 16 53.3 125 2 06DLX5_TENMO 06DLX5 tenebrio mo
994 16 53.3 125 2 0814U7_CAEBL 0814U7 caenorhabdi
995 16 53.3 125 2 038346_BPLIH 038346 laetococcus
996 16 53.3 125 2 04NR72_9DEL7 04NR72 anaeromyxob
997 16 53.3 125 2 08JFJ4_GINCI 08JFJ4 ginglymasto
998 16 53.3 125 2 08JFJ7_GINCI 08JFJ7 ginglymasto
999 16 53.3 125 2 08JFJ8_GINCI 08JFJ8 ginglymasto
1000 16 53.3 125 2 08JFJ0_GINCI 08JFJ0 ginglymasto
```

ALIGNMENTS

```
RESULT 1
Q9N669_9CAEN PRELIMINARY; PRT; 69 AA.
ID Q9N669;
AC Q9N669;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Four-loop conotoxin (Fragment).
OS Conus edraeus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxId=89425;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20416305; PubMed=10958845;
RA Duda T.F., Jr., Palumbi S.R.;
RT "Evolutionary diversification of multigene families: allelic selection
of toxins in predatory cone snails."
RL Mol. Biol. Evol. 17:1286-1293(2000).
RL [2]
RN NUCLEOTIDE SEQUENCE.
RP Duda T.F., Palumbi S.R.;
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF174281; AAF89945.1; -, mRNA.
DR EMBL; AF174280; AAF89944.1; -, mRNA.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0008200; F:ion channel inhibitor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR004214; Conotoxin.
DR Pfam; PF02950; Conotoxin; 1.
FT NON_TER 1
SQ SEQUENCE 69 AA; 7662 MW; CEC826FDC920C4F1 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 69;
Best local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Four-loop conotoxin (Fragment).
OS Conus edraeus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxId=89425;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20416305; PubMed=10958845;
RA Duda T.F., Jr., Palumbi S.R.;
RT "Evolutionary diversification of multigene families: allelic selection
of toxins in predatory cone snails."
RL Mol. Biol. Evol. 17:1286-1293(2000).
RL [2]
RN NUCLEOTIDE SEQUENCE.
RP Duda T.F., Palumbi S.R.;
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF174287; AAF89951.1; -, mRNA.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0008200; F:ion channel inhibitor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR004214; Conotoxin.
DR Pfam; PF02950; Conotoxin; 1.
FT NON_TER 1
SQ SEQUENCE 69 AA; 7674 MW; 83D526FDC934422A CRC64;
```

```
Query Match 60.0%; Score 18; DB 2; Length 69;
Best local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 4 CXXXXXXC 11
DB 61 CSTATSTC 68
```

```
RESULT 3
Q9NCS9_9CAEN PRELIMINARY; PRT; 69 AA.
ID Q9NCS9;
AC Q9NCS9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Four-loop conotoxin (Fragment).
OS Conus edraeus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxId=89425;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20416305; PubMed=10958845;
RA Duda T.F., Jr., Palumbi S.R.;
RT "Evolutionary diversification of multigene families: allelic selection
of toxins in predatory cone snails."
RL Mol. Biol. Evol. 17:1286-1293(2000).
RL [2]
RN NUCLEOTIDE SEQUENCE.
RP Duda T.F., Palumbi S.R.;
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF174286; AAF89950.1; -, mRNA.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0008200; F:ion channel inhibitor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR004214; Conotoxin.
DR Pfam; PF02950; Conotoxin; 1.
FT NON_TER 1
SQ SEQUENCE 69 AA; 7650 MW; DF3826FDC920C4F1 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 69;
Best local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

Qy 4 CXXXXXXC 11
|
Db 61 CSTATSTC 68

RESULT 4

Q9NCT2_9CAEN
ID Q9NCT2_9CAEN PRELIMINARY; PRT; 69 AA.
AC Q9NCT2;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Four-loop conotoxin (Fragment).
OS Conus ebraeus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Neogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=89425;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20416305; PubMed=10958845;
RA Duda T.F., Jr., Palumbi S.R.,
RT "Evolutionary diversification of multigene families: allelic selection
of toxins in predatory cone snails.";
RL Mol. Biol. Evol. 17:1286-1293(2000).
RN [2]

NUCLEOTIDE SEQUENCE.

RA Duda T.F., Palumbi S.R.,
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF174282; AAF8946.1; -, mRNA.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0008200; F:ion channel inhibitor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR004214; Conotoxin.
DR Pfam; PF02950; Conotoxin; 1.
FT NON_TER
SQ SEQUENCE 69 AA; 7634 MW; 463826FDC93E0BCD CRC64;

Query Match 60.0%; Score 18; DB 2; Length 69;
Best Local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11
|
Db 61 CSTATSTC 68

RESULT 5

Q9NCT3_9CAEN
ID Q9NCT3_9CAEN PRELIMINARY; PRT; 69 AA.
AC Q9NCT3;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Four-loop conotoxin (Fragment).
OS Conus ebraeus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=89425;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20416305; PubMed=10958845;
RA Duda T.F., Jr., Palumbi S.R.,
RT "Evolutionary diversification of multigene families: allelic selection
of toxins in predatory cone snails.";
RL Mol. Biol. Evol. 17:1286-1293(2000).
RN [2]

NUCLEOTIDE SEQUENCE.

RA Duda T.F., Palumbi S.R.,
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF174279; AAF89943.1; -, mRNA.

DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0008200; F:ion channel inhibitor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR004214; Conotoxin.
DR Pfam; PF02950; Conotoxin; 1.
FT NON_TER
SQ SEQUENCE 69 AA; 7643 MW; D5699CFDC920CAF1 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 69;
Best Local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11
|
Db 61 CSTATSTC 68

RESULT 6

Q9VD36_DROME
ID Q9VD36_DROME PRELIMINARY; PRT; 69 AA.
AC Q9VD36;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE CG13858-PA.
GN Name=CG13858; ORFNames=CG13858;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

NUCLEOTIDE SEQUENCE.

RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gockayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davignon L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Martel B., McIntosh T.C., McLeod J.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitakhs R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]

```

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celinker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champagne M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodegryn E.J.,
RA Svirskaas R., Taber P.B., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.B., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskaas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celinker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomic perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Miara S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,
RA Bellencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX Berkeley Drosophila Genome Project;
RA Celinker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskaas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RP NUCLEOTIDE SEQUENCE.
RX Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RA EMBL; AE003738; AAF55967.1; -; Genomic DNA.
DR Ensembl; CG13858; Drosophila melanogaster.
DR FlyBase; FBgn0040585; CG13858.
SQ SEQUENCE 69 AA; 6492 MW; EB077B26FD260BCC CRC64;

Query Match 60.0%; Score 18; DB 2; Length 69;
Best Local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 9 CSASSC 16

RESULT 7
Q70FP9_ANOGA PRELIMINARY; PRT; 90 AA.
AC Q70FP9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP00000019154 (Fragment).
GN ORFNames=ENSANG0000001665;
OS Anopheles gambiae str. PST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
NCBI_TaxID=180454;

```

```

RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB0100846; EAA06382.2; -; Genomic DNA.
FT NON TER 90
SQ SEQUENCE 90 AA; 9710 MW; 72DB579BDP140ADF CRC64;

Query Match 60.0%; Score 18; DB 2; Length 90;
Best Local Similarity 25.0%; Pred. No. 2.2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 20 CSASSAAC 27

RESULT 8
Q9B1I1_ANOGA PRELIMINARY; PRT; 90 AA.
AC Q9B1I1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE GSG1A protein precursor (Fragment).
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GASUA; TISSUE=Salivary glands;
RX MEDLINE=22057806; PubMed=12062411; DOI=10.1016/S0014-5793(02)02578-4;
RA Lanfranco A., Lombardo F., Santolamazza F., Veneri M.,
RA Castiglione T., Coluzzi M., Arca B.;
RT "Novel CDNA's encoding salivary proteins from the malaria vector
RT Anopheles gambiae.";
RL FEBS Lett. 517:67-71(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GASUA; TISSUE=Salivary glands;
RA Arca B.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ297932; CAC35452.1; -; mRNA.
DR Ensembl; ENSANG0000001665; Anopheles gambiae.
KW Signal.
FT SIGNAL 1 25 Potential;
FT NON TER 90 90
SQ SEQUENCE 90 AA; 9710 MW; 72DB579BDP140ADF CRC64;

Query Match 60.0%; Score 18; DB 2; Length 90;
Best Local Similarity 25.0%; Pred. No. 2.2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 20 CSASSAAC 27

RESULT 9
Q82854_9RETR PRELIMINARY; PRT; 114 AA.

```

AC 082854;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE TAR protein (Transactivating regulatory protein).
 GN Name=tat;
 OS Jembrana disease virus.
 OS Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
 OC Lentivirinae; Bovine lentiviruses.
 NC NCBI_TaxID=36370;
 RX MEDLINE=97201584; PubMed=9049370;
 RX Chadwick B.J., Coelen R.J., Wilcox G.E., Samuels L.M., Kertayadnya G.;
 "Nucleotide sequence analysis of Jembrana disease virus: a bovine
 lentivirus associated with an acute disease syndrome.";
 J. Gen. Virol. 76:1637-1650 (1995).
 RL EMBL; U21603; AAA64395.1; -; Genomic RNA.
 DR GO:0042025; C:host cell nucleus; IEA.
 DR GO:0005634; C:nucleus; IEA.
 DR GO:0003723; F:RNA binding; IEA.
 DR GO:0003700; F:transcription factor activity; IEA.
 DR GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO:0006350; P:transcription; IEA.
 DR InterPro: IPR001831; IV_Tat.
 DR Pfam: PF00539; Tat; 1.
 KW Activator; Nuclear protein; RNA-binding; Transcription;
 KW Transcription regulation.
 SQ SEQUENCE 114 AA; 12457 MW; 77A7C6B6FAF128D5A CRC64;

Query Match 60.0%; Score 18; DB 2; Length 114;
 Best Local Similarity 25.0%; Pred. No. 2.5e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
 DB 96 CASSASC 103

RESULT 10
 Q9EOL8 MOUSE PRELIMINARY; PRT; 161 AA.
 ID Q9EOL8; AC Q9EOL8;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE TIS1D deletion variant (Fragment).
 GN Name=zfp3612; Synonyms=Brf2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 NC NCBI_TaxID=10090;
 RX STRAIN=C57BLKS/J;
 RA Cho K., Hobson K., Greenhalgh D.G.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF205222; AAG45250.1; -; mRNA.
 DR HSSP; P22893; IM90.
 DR SMR; Q9EOL8; 1-49.
 DR MGI; MGI:107945; zfp3612.
 DR GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro: IPR000571; Znf_CCCH.
 DR Pfam; PF00642; zf-CCCH; 1.
 DR SMART; SM00356; znf_C3H1; 1.
 FT NON_TER 1 1
 FT SEQUENCE 161 AA; 17101 MW; BF09DF73709215EE CRC64;

Query Match 60.0%; Score 18; DB 2; Length 161;
 Best Local Similarity 25.0%; Pred. No. 2.8e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 11
 DB 129 CASSASC 136

RESULT 11
 Q7QPM4 GIALA PRELIMINARY; PRT; 163 AA.
 ID Q7QPM4; AC Q7QPM4;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE GIP 54 16521 16030.
 OS Giardia lamblia ATCC 50803.
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
 NC NCBI_TaxID=184922;
 RX NUCLEOTIDE SEQUENCE.
 RP STRAIN=WB C6;
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
 RA Olsen G.J., Sogin M.L.;
 RT "Draft sequence of the Giardia lamblia genome."
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 DR EMBL; AACB0100166; EAA36969.1; -; Genomic DNA.
 SQ SEQUENCE 163 AA; 17478 MW; 3AF4A5F8EDA9A10C CRC64;

Query Match 60.0%; Score 18; DB 2; Length 163;
 Best Local Similarity 25.0%; Pred. No. 2.8e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
 DB 24 CAATTAAC 31

RESULT 12
 Q8LDT9 ARATH PRELIMINARY; PRT; 177 AA.
 ID Q8LDT9; AC Q8LDT9;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eurosid 1; Brassicales; Brassicaceae; Arabidopsis.
 NC NCBI_TaxID=3702;
 RX MEDLINE=22088475; PubMed=12093376;
 RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 annotation.";
 RL Genome Biol. 3:RESEARCH0029-RESEARCH0029 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY085804; AAM63020.1; -; mRNA.
 DR Hypothetical protein.
 SQ SEQUENCE 177 AA; 19819 MW; CC6A7841CBA8B38 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 177;
 Best Local Similarity 25.0%; Pred. No. 2.9e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

Qy      4 CXXXXXXC 11
      |
Db      87 CXXXXSSC 94

RESULT 13
Q9C9L7_ARATH PRELIMINARY; PRT; 177 AA.
ID 09C9L7
AC 09C9L7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein T6C23.4.
GN Name=T6C23.4;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsia.
OX NCBI_TaxID=3702;
[1]
RN NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Renning C.M., Koo H., Fujii C.Y., Utterback T.R., C.M.;
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Town C.D., Kaul S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC013289; ANG52545.1; -; Genomic_DNA.
DR F1R; F96719; F96719.
KW Hypothetical protein.
SQ SEQUENCE 177 AA; 19834 MW; 03D870608FEB8246 CRC64;

Query Match      60.0%; Score 18; DB 2; Length 177;
Best Local Similarity 25.0%; Pred. No. 2.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy      4 CXXXXXXC 11
      |
Db      87 CXXXXSSC 94

RESULT 14
Q60M47_CABR PRELIMINARY; PRT; 199 AA.
ID 060M47
AC 060M47;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG23307.
GN Name=CBG23307;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderae; Caenorhabditis.
OX NCBI_TaxID=6238;
[1]
RN NUCLEOTIDE SEQUENCE.
RA The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 BPTI/Kunitz inhibitor domains.
DR EMBL; CAAC01000141; CAB75329.1; -; Genomic_DNA.
DR GO; GO:0004867; F:serine-type endopeptidase inhibitor activity; IEA.
DR InterPro; IPR002223; Prot inh_kunz-m.
DR InterPro; IPR006150; Worm repeat_1.
DR Pfam; PF00014; Kunitz_BPTI; 2.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Prot inh_kunz-m; 2.
DR SMART; SM00131; KU; 2.
DR SMART; SM00289; WRI; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS0279; BPTI_KUNITZ_2; 2.

```

```

KW Hypothetical protein.
SQ SEQUENCE 199 AA; 21668 MW; 380A14BB90A152A3 CRC64;

Query Match      60.0%; Score 18; DB 2; Length 199;
Best Local Similarity 25.0%; Pred. No. 3.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy      4 CXXXXXXC 11
      |
Db      112 CXXXXTTC 119

RESULT 15
Q5VMP6_ORISA PRELIMINARY; PRT; 200 AA.
ID 05VMP6
AC 05VMP6;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein OSUNBD008D07.38.
GN Name=OSUNBD008D07.38;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
[1]
RN NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arica K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikono M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karsawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshitara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006237; BAD69279.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 200 AA; 23098 MW; 0E4249F62A3678AA CRC64;

Query Match      60.0%; Score 18; DB 2; Length 200;
Best Local Similarity 25.0%; Pred. No. 3.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy      4 CXXXXXXC 11
      |
Db      102 CXXXXTTC 109

RESULT 16
Q9EOL7_MOUSE PRELIMINARY; PRT; 224 AA.
ID 09EOL7
AC 09EOL7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE TIS11D insertion variant (Fragment).
GN Name=Zfp3612; Synonym=Brf12;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

```


OC Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J;
 RA Cho K., Hobson K., Greenhalgh D.G.,
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF205223; AAC45251.1; -, mRNA.
 DR HSSP; P22893; 1M90.
 DR SMR; Q9EOL7; 1-49.
 DR MGI; MGI:107945; Zfp3612.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR000571; Znf_CCH.
 DR Pfam; PF00642; zfc_CCH; I.
 DR SMART; SM00356; ZNF_C3H1; 1.
 FT NON_TER 1
 FT NON_TER 224
 SQ SEQUENCE 224 AA; 23018 MW; 26E09C8465A5A61E CRC64;
 Query Match 60.0%; Score 18; DB 2; Length 224;
 Best Local Similarity 25.0%; Pred. No. 3.2e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 11
 DB 129 CASSASSC 136
 RESULT 17
 Q6K2G5_ORYSA
 ID Q6K2G5_ORYSA PRELIMINARY; PRT; 233 AA.
 AC Q6K2G5;
 DT 05-JUL-2004 (T-EMBLrel. 27, Created)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
 DE Glucanase-like.
 GN Name=OSJNB0052M16.18;
 OS *Oryza sativa* (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 2, BAC
 RT clone:OSJNB0052M16."
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005841; BAD23651.1; -, Genomic DNA.
 DR Gramene; Q6K2G5; -.
 SQ SEQUENCE 233 AA; 24907 MW; E4805DBE35FD0235 CRC64;
 Query Match 60.0%; Score 18; DB 2; Length 233;
 Best Local Similarity 25.0%; Pred. No. 3.3e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 11
 DB 199 CAATTAAC 206
 RESULT 18
 OS1XD3_MAGR
 ID OS1XD3_MAGR PRELIMINARY; PRT; 242 AA.
 AC OS1XD3;
 DT 13-SEP-2005 (T-EMBLrel. 31, Created)
 DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
 DE Predicted protein.
 GN ORFNames=MG07952.4;
 OS Magnaporthe grisea 70-15.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.

OX NCBI_TaxID=242507;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
 RA Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
 RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
 RA Bayul T., Bittsteyn B., Bloom T., Blye J., Boguslavsky L.,
 RA Borowsky M., Boukhalter B., Brunache A., Butler J., Calixte N.,
 RA Calvo S., Camarata J., Campo K., Chang J., Chesnatsang Y., Citroen M.,
 RA Collymore A., Consideine T., Cook A., Cooke P., Corum B., Cuomo C.,
 RA David R., Dawoe T., Degrey S., Dodge S., Doolley K., Dorje P.,
 RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
 RA Erickson J., Farina F., Faro S., Ferreira P., Fischer H.,
 RA Fitzgerald M., Foley K., Gage D., Galagan J., Garin G., Gierre S.,
 RA Ghitre A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafer N.,
 RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
 RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
 RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysseles M., Karlsson B.,
 RA Kells C., Kieu A., Klesner P., Kodira C., Kulbokas E., Labutti K.,
 RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
 RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
 RA Lui A., Ma L.J., Mabbitt R., MacDonald J., Maclean C., Major J.,
 RA Manning J., Marabelli R., Marr K., Matthews C., Mauceli E.,
 RA McCarthy M., McDonough S., Mcghee T., Meldrum J., Menues L.,
 RA Mesirov J., Mhalley A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
 RA Mozes J., Mulrain L., Munson G., Naylor J., Neves C., Nguyen C.,
 RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
 RA Norbu N., O'donnell P., Okawo O., O'leary S., Omotocho B.,
 RA O'Neill K., Osman S., Parker S., Perith D., Phunkhang P., Pigani B.,
 RA Purcell S., Rachupka T., Ramasamy U., Raneau R., Ray V., Raymond C.,
 RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
 RA Rutman M., Schubach R., Seaman C., Settipalli S., Sharpe T.,
 RA Sheridan J., Sherna N., Shi J., Smirnov S., Smith C., Sounguez C.,
 RA Spencer B., Stalker J., Strange-Thomann N., Scavopoulos S.,
 RA Stenson K., Stone S., Stone S., Stubbs M., Talamas J., Tchinga P.,
 RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
 RA Towey S., Tsamila T., Tsomo N., Vallée D., Vassiliev H.,
 RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
 RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
 RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembek L.,
 RA Zimmer A., Zody M., Zander E.;
 RT "The genome sequence of Magnaporthe grisea."
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AA001000977; EAA53675.1; -, Genomic DNA.
 SQ SEQUENCE 242 AA; 25532 MW; 3393ABE816D10BA3 CRC64;
 Query Match 60.0%; Score 18; DB 2; Length 242;
 Best Local Similarity 25.0%; Pred. No. 3.3e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 11
 DB 62 CASSSSC 69
 RESULT 19
 Q6BCP3_HPBVO
 ID Q6BCP3_HPBVO PRELIMINARY; PRT; 245 AA.
 AC Q6BCP3;

RA Tallon L.J., Feldblum T.V., Tsalir T., Bera J.J., Kim M.M., Jin S.,
 RA Padrosch D., Vuong H., Overton II L.L., Reardon M., Weaver B.,
 RA Johri S., Utecherak T.R., Pai G., Smith S., Mortman J., Haas B.J.,
 RA Zhu W., Yang Q., Koo H., Zismann V., Hsiao J., Quackenbush J.,
 RA White O., Salzberg S.L., Fraser C.M.,
 RT "Oryza sativa chromosome 3 BAC OSJNB0036M02 genomic sequence.";
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell R.;
 RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; ACL33930; AAP4639.1; -; Genomic_DNA.
 DR EMBL; ACL45388; AAV089142.1; -; Genomic_DNA.
 DR Gramene; Q7Y198; -;
 DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0016567; P:protein ubiquitination; IEA.
 DR InterPro; IPR001841; Znf_ring.
 DR SMART; SM00184; RING; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 319 AA; 33316 MW; FD51B063E2B4AB4 CRC64;
 Query Match 60.0%; Score 18; DB 2; Length 319;
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 11
 DB 296 CAAAAAAC 303

RESULT 23
 O61A5 CAEBR PRELIMINARY; PRT; 320 AA.
 ID O61A5;
 AC O61A5;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, last annotation update)
 DE Hypothetical protein CBG13750.
 GN Name=CBG13750;
 OS Caenorhabditis briggsae.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6238;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RG The C.briggsae Sequencing Consortium;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; CAC0100066; CAE68107.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004930; F:G-protein coupled receptor activity; IEA.
 DR InterPro; IPR000168; Nm7TM_chemrecept.
 KW Hypothetical protein.
 SQ SEQUENCE 320 AA; 36037 MW; 5E3389F6C7CF3C7B CRC64;
 Query Match 60.0%; Score 18; DB 2; Length 320;
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 11
 DB 302 CASSTSTC 309

RESULT 24
 O6ZNM3 HUMAN PRELIMINARY; PRT; 328 AA.
 ID O6ZNM3;
 AC O6ZNM3;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, last annotation update)
 DE Hypothetical protein FLJ29006.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Stomach mucosa;
 RA Nimmla K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
 RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
 RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Suzuki Y., Hata H.,
 RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,
 RA Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,
 RA Nagai K., Isegai T., Sugano S.,
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK131040; BAC85481.1; -; mRNA.
 KW SEQUENCE 328 AA; 35381 MW; 7D04FC08E3270F8D CRC64;
 Query Match 60.0%; Score 18; DB 2; Length 328;
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 11
 DB 4 CTSASSSC 11

RESULT 25
 O6ZAB6 ORYZA PRELIMINARY; PRT; 332 AA.
 ID O6ZAB6;
 AC O6ZAB6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, last annotation update)
 DE Limonene cyclase like protein.
 GN Name=OSJNB0084107.20-2; Synonyms=OSJNB0002109.8-2;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 RT clone:OSJNB0084107.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 RT clone:OSJNB0002109.";
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005179; BAC83914.1; -; Genomic_DNA.
 DR EMBL; AP005877; BAC31847.1; -; Genomic_DNA.
 DR Gramene; Q6Z486; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR InterPro; IPR004263; Exostosin.
 DR Pfam; PF03016; Exostosin; 1.
 KW SEQUENCE 332 AA; 34451 MW; 57CD6BBFC0CC5D6E CRC64;
 Query Match 60.0%; Score 18; DB 2; Length 332;
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 11
 DB 27 CAAAAAAC 34

```

RESULT 26
Q94589_9CILI PRELIMINARY; PRT; 350 AA.
ID Q94589_9CILI
AC Q94589
DT 01-FEB-1997 (TREMblrel. 02, Created)
DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Lemnaden factor.
OS Lemnaden bullinum.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
OC Lemnaden.
OX NCBI_TaxID=54108;
RX NUCLEOTIDE SEQUENCE.
RP Peters-Regehr T., Kusch J., Heckmann K.;
RA Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL; Y09220; CAA70420.1; -, Genomic_DNA.
DR HSSP; O16119; 1E2G.
DR InterPro; IPR006212; Furin_repeat.
DR SMART; SM00261; FU; 3.
SQ SEQUENCE 350 AA; 35159 MW; DBE0C67654B9D92E CRC64;

Query Match 60.0%; Score 18; DB 2; Length 350;
Best Local Similarity 25.0%; Pred. No. 3.8e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11
Db 105 CSTSATTC 112

RESULT 27
RHOM DROME STANDARD; PRT; 355 AA.
ID P20350; Q9W0F2;
AC 01-FEB-1991 (Rel. 17, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Rhomboid protein (EC 3.4.21.-) (Veinlet protein).
GN Name=rho; Synonyms=Ve; ORFNames=CG1004;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriodes; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RX NUCLEOTIDE SEQUENCE.
RP MEDLINE=90249726; PubMed=2110920;
RA Bier E., Jan L.Y., Jan Y.N.;
RT "Rhomboid, a gene required for dorsoventral axis establishment and
peripheral nervous system development in Drosophila melanogaster.";
RL Genes Dev. 4:190-203 (1990).
[2]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abiri J.F., Agbayani A., An H.-O., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borokova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Cantor A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mayes A.D., Dew I., Dietz S.M.,
Doudan K.J., Dougan M., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

```

```

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murzyn D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach U.,
RA Williams S.M., Woodedge T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
[3]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Miura S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Betencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).
[4]
RP FUNCTION, AND MUTAGENESIS OF TRP-151; ARG-152; ASN-169; GLY-215;
RP SER-217 AND HIS-281.
RX MEDLINE=21526629; PubMed=11672525; DOI=10.1016/S0092-8674(01)00525-6;
RA Urban S., Lee J.R., Freeman M.;
RT "Drosophila Rhomboid-1 defines a family of putative intramembrane
serine proteases.";
RL Cell 107:173-182 (2001).
CC -!- FUNCTION: Acts early in embryonic development to establish
position along the dorsoventral axis and then again later to
specify the fate of neuronal precursor cells. Involved in EGF
receptor signaling; cleaves Spitz to release the active growth
factor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Golgi.
CC -!- DEVELOPMENTAL STAGE: Early blastoderm stages and later during
neuron development.
CC -!- SIMILARITY: Belongs to the peptidase S54 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
EMBL; X52454; CAA36692.1; -, mRNA.
EMBL; AE003471; AAF47496.1; -, Genomic_DNA.
DR PIR; A34597; A34597.
DR MEROPS; S54.001; -.
DR Ensembl; CG1004; Drosophila melanogaster.
DR Flybase; FBgn0004635; rho.
DR GO; GO:0005794; C:Golgi apparatus; IDA.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0008236; P:serine-type peptidase activity; IDA.
DR GO; GO:0007420; P:brain development; IMP.
DR GO; GO:0046445; P:branched duct epithelial cell fate determin. .; TAS.
DR GO; GO:0001763; P:branching morphogenesis; NAS.
DR GO; GO:0007174; P:epidermal growth factor ligand processing; IGI.
DR GO; GO:0007479; P:leg disc proximal/distal pattern formation; TAS.
DR GO; GO:0007438; P:neococyte development; IGI.

```

DR GO; GO:0030707; P:ovarian follicle cell development (sensu In. . .; TAS.
 DR GO; GO:0007422; P:peripheral nervous system development; TAS.
 DR GO; GO:0045742; P:positive regulation of epidermal growth fac. . .; TAS.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IGI.
 DR GO; GO:0007176; P:regulation of epidermal growth factor recep. . .; TAS.
 DR GO; GO:0007432; P:salivary gland determination; NAS.
 DR GO; GO:0005502; P:tracheal sac formation (sensu Insecta); TAS.
 DR GO; GO:0007474; P:wing vein specification; NAS.
 DR InterPro:IPR002610; Rhomboid_Like.
 DR Pfam:PF01694; Rhomboid; 1.
 KW Developmental protein; Golgi stack; Hydrolase; Protease;
 KW Serine protease; Transmembrane.
 FT TOPO_DOM 1 98
 FT TRANSMEM 99 119
 FT TOPO_DOM 120 162
 FT TRANSMEM 163 183
 FT TOPO_DOM 184 188
 FT TRANSMEM 189 209
 FT TOPO_DOM 210 210
 FT TRANSMEM 211 231
 FT TOPO_DOM 232 244
 FT TRANSMEM 245 265
 FT TOPO_DOM 266 275
 FT TRANSMEM 276 296
 FT TOPO_DOM 297 308
 FT TRANSMEM 309 329
 FT TOPO_DOM 330 355
 FT ACT_SITE 169 169
 FT ACT_SITE 217 217
 FT ACT_SITE 281 281
 FT MUTAGEN 151 151
 FT MUTAGEN 152 152
 FT MUTAGEN 169 169
 FT MUTAGEN 215 215
 FT MUTAGEN 217 217
 FT MUTAGEN 281 281
 FT CONFLICT 4 4
 FT CONFLICT 33 33
 FT CONFLICT 46 46
 FT CONFLICT 46 46
 SQ SEQUENCE 355 AA; 39330 MW; 2E0572D08831C5D0 CRC64;
 Query Match 60.0%; Score 18; DB 1; Length 355;
 Best Local Similarity 25.0%; Pred. No. 3.8e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 11
 DB 57 CSTASSTC 64
 RESULT 28
 Q540V7 DROME
 ID Q540V7; DROME PRELIMINARY; PRT; 355 AA.
 AC Q540V7;
 DT 13-SRP-2005 (T-EMBLrel. 31, Created)
 DT 13-SRP-2005 (T-EMBLrel. 31, Last sequence update)
 DT 13-SRP-2005 (T-EMBLrel. 31, Last annotation update)
 DE ID06131P.
 GN Name=ro;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_Taxid=7227;
 RX NUCLEOTIDE SEQUENCE.
 RC STRAIN=Berkeley.
 RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nuno U., Pacle U., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY121641; AAM51968.1; -, mRNA.
 SQ SEQUENCE 355 AA; 39330 MW; 2E0572D08831C5D0 CRC64;
 Query Match 60.0%; Score 18; DB 2; Length 355;
 Best Local Similarity 25.0%; Pred. No. 3.8e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 11
 DB 57 CSTASSTC 64
 RESULT 29
 T1SD MOUSE
 ID T1SD MOUSE STANDARD; PRT; 367 AA.
 AC P23949;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Butyrate response factor 2 (T1SID protein).
 GN Name-zfp3612; Synonyms=Brf2, T1SID;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OC NCBI_Taxid=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=BALB/c;
 RA MEDLINE=91141531; PubMed=1996120;
 RA Varnum B.C., Ma Q., Chi T., Fletcher B., Herschman H.R.;
 RT "The T1SID primary response gene is a member of a gene family that
 RT encodes proteins with a highly conserved sequence containing an
 RT unusual Gys-His repeat.";
 RL Mol. Cell. Biol. 11:1754-1758 (1991).
 RN [2]
 RP NUCLEOTIDE SEQUENCE OF 1-50.
 RA Fletcher B.S.;
 RL Thesis (1992); University of California Los Angeles, United States.
 CC -1- FUNCTION: Probable regulatory protein involved in regulating the
 CC response to growth factors.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- SIMILARITY: Contains 2 C3H1-type zinc fingers.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC EMBL; M58564; AAA72946.1; -, mRNA.
 DR EMBL; M57165; AAA9709.1; -, Genomic DNA.
 DR PIR; C39590; C39590.
 DR HSSP; P22893; 1M90.
 DR SMR; P23949; 124-193.
 DR Ensembl; ENSMUSG0000045817; Mus musculus.
 DR MGI; MGI:107945; Zfp3612.
 DR InterPro; IPR007635; T1SID_N.
 DR InterPro; IPR000571; ZnF_CCH.
 DR Pfam; PF04553; T1SID_C; 1.
 DR Pfam; PF00642; zf-CCH; 2.
 DR SMART; SM00356; ZnF_C3H1; 2.
 KW DNA-binding; Metal-Binding; Nuclear protein; Repeat; Zinc;
 KW Zinc-finger.
 FT ZN_FING 132 151
 FT ZN_FING 170 189
 FT COMPBIAS 64 67
 FT COMPBIAS 111 114
 FT COMPBIAS 200 203
 FT COMPBIAS 263 266
 FT COMPBIAS 297 303
 SQ SEQUENCE 367 AA; 37593 MW; 361244AF6244E46E CRC64;

Query Match 60.0%; Score 18; DB 1; Length 367;
 Best Local Similarity 25.0%; Pred. No. 3.9e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
 DB 273 CSSSSSSC 280

RESULT 30

Q66SB6_DROVI PRELIMINARY; PRT; 404 AA.
 AC Q66SB6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DE Rhoimoid.
 GN Name=rho.
 OS Drosophila virilis (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OK NCBI_TaxID=7244;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22844358; PubMed=12963110; DOI=10.1016/S0925-4773(03)00164-3;
 RA Nakamura Y., Matsuno K.;
 RT "Species-specific activation of EGF receptor signaling underlies
 RT evolutionary diversity in the dorsal appendage number of the genus
 RT Drosophila eggshells.";
 RL Mech. Dev. 120:897-907(2003).
 DR EMBL; AB089248; BAC56701.1; -; mRNA.
 DR MEROPS; S54.001; -;
 DR FLYBASE; FBgn0062278; Dvir/rho.
 DR InterPro; IPR002610; Rhoimoid_Like.
 DR Pfam; PF01694; Rhoimoid; 1.
 SQ SEQUENCE 404 AA; 44035 MW; EF641632B02F9011 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 404;
 Best Local Similarity 25.0%; Pred. No. 4e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
 DB 63 CTFASSTC 70

RESULT 31

Q624B5_ORYSA PRELIMINARY; PRT; 417 AA.
 AC Q624B5;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Limonene cyclase like protein.
 GN Name=OSUNB0084L07.20-3; Synonyms=OSUNB0002L09.8-3;
 OS Oryza sativa (Japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoidae; Oryzaceae; Oryza.
 OK NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 RT clone:OSUNB0084L07.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 RT clone:OSUNB0002L09.";

RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005179; BAC83915.1; -; Genomic_DNA.
 DR EMBL; AP005877; BAC31849.1; -; Genomic_DNA.
 DR Gramene; Q624B5; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR InterPro; IPR004263; Exostein.
 DR Pfam; PF03016; Exostein; 1.
 SQ SEQUENCE 417 AA; 44610 MW; 67668B9A53B1DC6A CRC64;

Query Match 60.0%; Score 18; DB 2; Length 417;
 Best Local Similarity 25.0%; Pred. No. 4.1e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
 DB 27 CAAAAAAC 34

RESULT 32

Q80UL4_ORYLA PRELIMINARY; PRT; 445 AA.
 AC Q80UL4;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DE TAPBP protein.
 GN Name=TAPBP.
 OS Oryzias latipes (Medaka fish) (Japanese ricefish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Atherinomorpha;
 OC Belontiiformes; Adriantichthyidae; Oryziatidae; Oryzias.
 OK NCBI_TaxID=8090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Hd-IR;
 RX MEDLINE=21850510; PubMed=11862394; DOI=10.1007/s00251-001-0427-3;
 RA Matsuo M.Y., Aakawa S., Shimizu N., Kimura H., Nonaka M.;
 RT "Nucleotide sequence of the MHC class I genomic region of a teleost,
 RT the medaka (Oryzias latipes).";
 RL Immunogenetics 53:930-940(2002).
 DR EMBL; BA000027; BAB83851.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0019885; P:antigen processing, endogenous antigen via . . .; IEA.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR008056; Tapasin.
 DR Pfam; PF00047; Ig; 1.
 DR PRINTS; PR01669; TAPASIN.
 DR SMART; SM00409; Ig; 2.
 DR PROSITE; PS50835; IG_LIKE; 2.
 DR KEGG; K01101; Immunoglobulin domain; Repeat.
 SQ SEQUENCE 445 AA; 48260 MW; FDF7C5FBE74370BB CRC64;

Query Match 60.0%; Score 18; DB 2; Length 445;
 Best Local Similarity 25.0%; Pred. No. 4.2e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
 DB 22 CSSSSSSC 29

RESULT 33

Q7SHC6_NEUCR PRELIMINARY; PRT; 480 AA.
 ID Q7SHC6_NEUCR PRELIMINARY;
 AC Q7SHC6;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DE Predicted protein (Hypothetical protein B13N4.210).
 GN Name=NCU01880.1; Synonyms=B13N4.210;
 OS Neurospora crassa.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 NCBI_TaxId=5141;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=0874A;
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
 RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
 RA Eklins T., Engels S., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
 RA Qui D., Tankler P., Pedersen D., Nelson M., Washburne M.,
 RA Selitrenikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
 RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
 RA Roy A., Foley K., Naylor J., Thomann N., Barcott R., Gnerre S.,
 RA Kraml M., Kamyshelev M., Mauceli E., Bielke C., Rudd S., Fishman D.,
 RA Kryzofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
 RA Cogoni C., Macino G., Catchside D., Il W., Pratt R.J., Osmani S.A.,
 RA Desouza C.C., Glaes L., Orbach M.J., Berglund J., Voelker R.,
 RA Yarden O., Plamann M., Sella S., Dunlap J., Radford A., Aramayo R.,
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbold D.J., Freitag M.,
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
 RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
 RL Nature 0:0-0(2003).
 RN [2]
 RN NUCLEOTIDE SEQUENCE.
 RA Schulte U., Aign V., Hohnsels J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN NUCLEOTIDE SEQUENCE.
 RA German Neurospora genome project;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AABX01000004; EAA36275.1; -; Genomic DNA.
 DR EMBL; BX842681; CAE81980.1; -; Genomic DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 480 AA; 50449 MW; E073DB8EC33C8A2B CRC64;
 QY 4 CXXXXXXC 11
 Db 127 CSSSSSSC 134
 Query Match 60.0%; Score 18; DB 2; Length 480;
 Best Local Similarity 25.0%; Pred. No. 4.3e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

DR GO; GO:0004114; F.3.'5'-cyclic-nucleotide phosphodiesterase a. .; IEA.
 DR GO; GO:0003824; Fcatalytic activity; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR003607; Met_phos_hydro.
 DR InterPro; IPR002073; PDEase.
 DR Pfam; PF00233; PDEase_I; 1.
 DR PRINTS; PR00387; PDIESTERASE1.
 DR SMART; SM00471; Hdc; 1.
 DR PROSITE; PS00126; PDEASE_I; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 480 AA; 53835 MW; 75AB3894A55D01CB CRC64;
 QY 4 CXXXXXXC 11
 Db 205 CTAAATAC 212
 Query Match 60.0%; Score 18; DB 2; Length 480;
 Best Local Similarity 25.0%; Pred. No. 4.3e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

RESULT 35
 T1SD_HUMAN
 ID T1SD_HUMAN STANDARD; PRT; 492 AA.
 AC P47974; Q9BSJ3;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Butyrate response factor 2 (T1SID protein) (EGF-response factor 2)
 DE (BRF-2).
 GN Name=ZFP3612; Synonyms=BRF2, ERF2, T1SID;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
 OC Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=96132724; PubMed=8545129;
 RA Ino T., Yasui H., Hirano M., Kurosawa Y.;
 RT "Identification of a member of the T1SID early response gene family at
 RT the insertion point of a DNA fragment containing a gene for the T-cell
 RT receptor beta chain in an acute T-cell leukemia.";
 RL Oncogene 11:2705-2710(1995).
 RN [2]
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=95137407; PubMed=7835719; DOI=10.1016/0378-1119(94)00696-P;
 RA Nle X.F., Maclean K.N., Kumar V., McKay I.A., Bustin S.A.;
 RT "ERF-2, the human homologue of the murine T1sid early response
 RT gene.";
 RL Gene 152:285-286(1995).
 RN [3]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RP TISSUE=Pancreas;
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Haich P.,
 RA Diatchenko L., Marusina B., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Umed T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Lotteliano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richardson S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.B.,
 RA Scherch A., Schein J.E., Jones S.J.W., Marra W.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human

```

RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Probable regulatory protein involved in regulating the
CC response to growth factors.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: Contains 2 C3H1-type zinc fingers.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
CC EMBL: U07802; AAA1778.1; -; Genomic_DNA.
DR EMBL: X78992; CAAS5592.1; -; mRNA.
DR EMBL: BC005010; AAH05010.1; -; mRNA.
DR PIR: S49147; S49147.
DR PDB: 1RGO; NMR: A=151-220.
DR Ensemble: ENSG00000152518; Homo sapiens.
DR HGNC: HGNC:1108; ZFP36L2.
DR H-InvDB: HIX0002007; -.
DR GO: GO:0003700; F:transcription factor activity; TAS.
DR GO: GO:0008283; P:cell proliferation; TAS.
DR InterPro: IPR007635; Tis11B_N.
DR InterPro: IPR000571; Znf_CCGH.
DR Pfam: PF04553; Tis11B_C; 1.
DR Pfam: PF00642; zf-CCGH; 2.
DR SMART: SM00356; ZNF_C3H1; 2.
DR 3D-structure: DNA-binding; Metal-binding; Nuclear protein; Repeat;
KW Zinc; zinc-finger.
FT ZN_FING 159 178 C3H1-type 1.
FT ZN_FING 197 216 C3H1-type 2.
FT COMPBIAS 106 109 Poly-Gly.
FT COMPBIAS 138 141 Poly-Gln.
FT COMPBIAS 143 146 Poly-Gly.
FT COMPBIAS 288 291 Poly-Pro.
FT COMPBIAS 323 330 Poly-Ala.
FT COMPBIAS 382 388 Poly-Ala.
FT COMPBIAS 393 397 Poly-Gln.
FT COMPBIAS 96 97 TS -> DL (in Ref. 1).
FT CONFLICT 318 318 A -> T (in Ref. 2).
FT CONFLICT 329 330 AA -> LR (in Ref. 2).
FT CONFLICT 330 330 A -> AAA (in Ref. 3).
FT CONFLICT 399 399 A -> QQQ (in Ref. 3).
FT CONFLICT 451 460 Missing (in Ref. 1).
SQ SEQUENCE 492 AA; 50921 MW; E4E2EE26791CCP7 CRC64;

Query Match 60.0%; Score 18; DB 1; Length 492;
Best Local Similarity 25.0%; Pred. No. 4.4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 299 CSSSASSC 306

RESULT 36
ID Q53TB4 HUMAN PRELIMINARY; PRT; 494 AA.
AC Q53TB4;
DT 13-SEP-2005 (TEMBLrel. 31, Created)
DT 13-SEP-2005 (TEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TEMBLrel. 31, Last annotation update)
DE Hypothetical protein ZFP36L2.
DE Name=ZFP36L2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OC NCB1 TaxID=9606;
RP NUCLEOTIDE SEQUENCE.
RP Abbot A., Boyer E., Heyen J.;

```

```

RT "The sequence of Homo sapiens BAC clone RP11-339H12."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RL Waterston R.H.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RL Waterston R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC010883; AA14992.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 494 AA; 51062 MW; 10E23FAE2DDABD4 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 494;
Best Local Similarity 25.0%; Pred. No. 4.4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 299 CSSSASSC 306

RESULT 37
ID Q6PD21 MOUSE PRELIMINARY; PRT; 526 AA.
AC Q6PD21;
DT 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE Shb protein (Fragment).
OS Name=Shb;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCB1 TaxID=10090;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek U.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Rahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Maria M.A.;
RT "generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC058986; AAH58986.1; -; mRNA.
DR HSSP: P00524; 1KC2.
DR GO: GO:0005515; F:protein binding; IDA.

```



```

DR GO; GO:0005070; F:SH3/SH2 adaptor activity; TAS.
DR GO; GO:0042100; P:B-cell proliferation; IDA.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
FT NON TER
SQ SEQUENCE 526 AA; 57393 MW; 6F358E391E2C63E0 CRC64;

Query Match
Best Local Similarity 25.0%; Score 18; DB 2; Length 526;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 71 CSASASAC 78

RESULT 38
Q5B195 DROME PRELIMINARY; PRT; 544 AA.
ID Q5B195 DROME PRELIMINARY; PRT; 544 AA.
AC Q5B195;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE RE04051P.
GN Name=dpa;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN (1)
RX NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkley;
RA Stapleton M., Carlson J., Chavez C., Frise E., George R., Paclieb J.,
RA Park S., Wan K., Yu C., Rubin G.M., Celinkier S.;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT021329; AAK3437.1; -; mRNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005924; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008094; F:DNA-dependent ATPase activity; IEA.
DR GO; GO:0006270; F:DNA replication initiation; IEA.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR001208; MCM.
DR InterPro; IPR008047; MCM 4.
DR PRINTS; PR01660; MCMPROTEIN4.
DR SMART; SM00350; MCM; 1.
DR PROSITE; PS00615; C:TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE; PS50051; MCM 2; 1.
SQ SEQUENCE 544 AA; 60347 MW; 8D221F998960D707 CRC64;

Query Match
Best Local Similarity 25.0%; Score 18; DB 2; Length 544;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 522 CSTRSTSC 529

RESULT 39
Q91TR2 TUHV1 PRELIMINARY; PRT; 545 AA.
ID Q91TR2 TUHV1 PRELIMINARY; PRT; 545 AA.
AC Q91TR2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE T31.
Tupaid herpesvirus 1 (strain 1) (TuHV-1) (Herpesvirus tupai (strain

```

```

OS 1)).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
NCBI_TaxID=10397;
RN (1)
RX NUCLEOTIDE SEQUENCE.
RC STRAIN=2;
RC MEDLINE=21211637; PubMed=11312357;
RX DOI=10.1126/JVI.75.10.4854-4870.2001;
RA Bahr U., Darai G.;
RT "Analysis and characterization of the complete genome of tupai (tree
RT shrew) herpesvirus.";
RL J. Virol. 75:4854-4870(2001).
DR EMBL; AF281817; AAK57075.1; -; Genomic_DNA.
DR InterPro; IPR007578; DUF570.
DR Pfam; PF04489; DUF570; 1.
SQ SEQUENCE 545 AA; 59508 MW; 98B5ECDEFF7A806 CRC64;

Query Match
Best Local Similarity 25.0%; Score 18; DB 2; Length 545;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 535 CAASATC 542

RESULT 40
Q528N9 ORYSA PRELIMINARY; PRT; 551 AA.
ID Q528N9 ORYSA PRELIMINARY; PRT; 551 AA.
AC Q528N9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Aspartic proteinase nepenthesin II-like.
GN Name=P0541C02.19-2;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
NCBI_TaxID=39947;
RN (1)
RX NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
RT clone:P0541C02.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003769; BAD61723.1; -; Genomic DNA.
DR GO; GO:0004194; F:pepsin A activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001461; Peptidase_A1.
DR InterPro; IPR001969; Pept_Asp_AS.
DR Pfam; PF00026; Asp; 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
SQ SEQUENCE 551 AA; 57586 MW; BE78840EB14824DF CRC64;

Query Match
Best Local Similarity 25.0%; Score 18; DB 2; Length 551;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 182 CATATSSC 189

RESULT 41
Q8SB30 ORYSA PRELIMINARY; PRT; 551 AA.
ID Q8SB30 ORYSA PRELIMINARY; PRT; 551 AA.
AC Q8SB30;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

```

DE Putative chloroplast nucleoid DNA-binding protein.
GN Name=OJ1540.H01.13;
OC Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriatroidae; Oryzaceae; Oryza.
OC NCBI_TaxID=4530;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan O., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Ganabinger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Teitelin T., Riggs F., Hsiao J., Zisman V., Blunt S., Pai G.,
RA Vanaken S.B., Utterback T.R., Feldblum T.V., Kalb E., Quackenbush J.,
RA Salzberg S.L., White O., Fraser C.M.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC091774; AAL79734.1; -; Genomic_DNA.
DR HSSP; P00797; 2REN.
DR Gramene; O8SB30; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004194; F:pepsin A activity; IEA.
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR01461; Peptidase_A1.
DR InterPro; IPR01969; Pept_Asp_AS.
DR Pfam; PF00026; Asp; 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
DR DNA-binding.
SQ SEQUENCE 551 AA; 57586 MW; BE78840EB14824DF CRC64;

Query Match 60.0%; Score 18; DB 2; Length 551;
Best Local Similarity 25.0%; Pred. No. 4.6e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 182 CATATSSC 189

RESULT 42
Q5AP57 CANAL PRELIMINARY; PRT; 552 AA.
AC Q5AP57;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical YFW family protein 5.
GN ORFNames=Ca019.4881;
OS Candida albicans SC5314.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OC NCBI_TaxID=337561;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=SC5314;
RX PubMed=15123810; DOI=10.1073/pnas.0401648101;
RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,
RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
RA Davis R.W., Scherer S.;
RT "The diploid genome sequence of Candida albicans."
RL Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=SC5314;
RA Dungan J., Kuo A., Newport G., Ian C.-Y., Iijima C., Adegbola O.,
RA Roberts J., Persson K., Donnelly S., Favoretto S., Tsung K.-W.,
RA Jones T., Scherer S., Agabian N.;
RT "Annotation of the Genome of Candida albicans."
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAC001000001; EAL04869.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 552 AA; 62256 MW; 5CA7280BD5B87B3F CRC64;

Query Match 60.0%; Score 18; DB 2; Length 552;
Best Local Similarity 25.0%; Pred. No. 4.6e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 466 CTSATTSC 473

RESULT 43
O80J53 HPBV0
ID O80J53 HPBV0 PRELIMINARY; PRT; 560 AA.
AC O80J53;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Truncated polymerase.
OC Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OC NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Xu X., Zheng D.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY238972; ANO64453.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003664; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0015740; F:transferase activity; IEA.
DR GO; GO:006278; F:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; RVT_1; 1.
SQ SEQUENCE 560 AA; 62448 MW; 7EC0D0F6F991CE4B CRC64;

Query Match 60.0%; Score 18; DB 2; Length 560;
Best Local Similarity 25.0%; Pred. No. 4.6e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 268 CASSSSSC 275

RESULT 44
Q6N1S1 CORDI
ID Q6N1S1 CORDI PRELIMINARY; PRT; 581 AA.
AC Q6N1S1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative secreted protein.
GN Ordered locus Names=D1P0696;
OC Corynebacterium diphtheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OC NCBI_TaxID=1717;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=Bloeype graves / NCTC 13129;
RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gk9874;
RA Cerdano-Tarraga A.-M., Estratou A., Dover L.G., Holden M.T.G.,

RA Pallen M.J., Bentley S.D., Beesa G.S., Churcher C.M., James K.D.,
 RA De Zoya A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
 RA Hamlin N., Holroyd S., Jagsels K., Moule S., Quail M.A.,
 RA Rabinowitsch E., Rutherford K.M., Thomson N.R., Unwin L.,
 RA Whitehead S., Barrett B.G., Parkhill J.,
 RT "The complete genome sequence and analysis of *Corynebacterium*
 RT *diphtheriae* NCTC13129.";
 RL Nucleic Acids Res. 31:6516-6523(2003).
 DR EMBL: BX248355; CA649213.1; -; Genomic_DNA.
 KW Complete proteome.
 SQ SEQUENCE 581 AA; 62851 MW; 9CF98B85A311AC3 CRC64;
 Query Match 60.0%; Score 18; DB 2; Length 581;
 Best Local Similarity 25.0%; Pred. No. 4.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 11
 DB 17 CAATTAAC 24
 RESULT 45
 Q51VZ0_MAGGR PRELIMINARY; PRT; 592 AA.
 ID Q51VZ0_MAGGR PRELIMINARY;
 AC Q51VZ0;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=MG03995.4;
 OS Magnaporthe oryzae 70-15.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetes; incertae sedis; Magnaportheaceae; Magnaporthe.
 OC NCBI_TaxID=242507;
 RN [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
 RA Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
 RA Arachchi H., Amburster J., Bachansang P., Baldwin J., Barry A.,
 RA Bayul T., Blithstein B., Bloom T., Bye J., Boguslavsky L.,
 RA Borowsky M., Boukhalter B., Brumache A., Butler J., Calixte N.,
 RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
 RA Collymore A., Considine T., Cook A., Cooke P., Cornum B., Cuomo C.,
 RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
 RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
 RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
 RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gierre S.,
 RA Gintre A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafer N.,
 RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
 RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Huebly E., Iliev I.,
 RA Jaffe D., Jones C., Kamal W., Kamat A., Kamysseis M., Karlsson E.,
 RA Kells C., Klen A., Kisaner P., Kodira C., Kulbokas E., Labutti K.,
 RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
 RA Lindblad-toh K., Liu X., Lokytang T., Lokytang Y., Lucien O.,
 RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
 RA Manning J., Marabella R., Maru K., Matthews C., Mancini E.,
 RA McCarty M., McDonough S., Meghee T., Meltrin J., Menais L.,
 RA Mesitov J., Mihalov A., Mihova T., Mikkelson T., Mianga V., Moru K.,
 RA Mozes J., Mulrain L., Munson G., Naylor J., Neues C., Nguyen C.,
 RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
 RA Norbu N., O'donnell P., Okawo O., O'leary S., Omotosho B.,
 RA O'Neill K., Osman S., Parker S., Perrin D., Phunhang P., Pigani B.,
 RA Purcell R., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
 RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
 RA Rutman M., Schipbach R., Seaman C., Settipalli S., Sharpe T.,
 RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Souguez C.,
 RA Spencer B., Stalker J., Stange-Thomann N., Stavropoulos S.,
 RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchinga P.,
 RA Tenzing P., Teefaye S., Theodore J., Thoulutang Y., Topham K.,
 RA Towey S., Tsamila T., Tsomo N., Vallie D., Vassiliev H.,
 RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
 RA Wangli T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,

RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
 RA Zimmer A., Zody M., Lander E.,
 RA "The genome sequence of *Magnaporthe oryzae*."
 RT Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Dean R., Mitchell T., Brown D., Pan H., Thon M.,
 RT Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.,
 RT Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL: AACU0100148; EAA50236.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 592 AA; 63027 MW; 84D30A8CA5D8ABC CRC64;
 Query Match 60.0%; Score 18; DB 2; Length 592;
 Best Local Similarity 25.0%; Pred. No. 4.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 11
 DB 347 CAAASTSC 354
 RESULT 46
 Q6Z4E7_ORYSA PRELIMINARY; PRT; 606 AA.
 ID Q6Z4E7_ORYSA PRELIMINARY;
 AC Q6Z4E7;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Putative pectin-glucuronyltransferase.
 GN Name=OSJNB0084L07.20-1; Synonyms=OSJNB0002L09.8-1;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=39947;
 RN [1]
 RN NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.,
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 RT clone:OSJNB0084L07.";
 RN Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.,
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 RT clone:OSJNB0084L07.";
 RN Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF005179; BAC83913.1; -; Genomic_DNA.
 DR EMBL: AF005877; BAC31848.1; -; Genomic_DNA.
 DR Gramene; Q6Z4E7; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR004263; Exostosin.
 DR Pfam; PF03016; Exostosin; 1.
 KW Transferase.
 SQ SEQUENCE 606 AA; 65830 MW; P90248B809A18266 CRC64;
 Query Match 60.0%; Score 18; DB 2; Length 606;
 Best Local Similarity 25.0%; Pred. No. 4.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 11
 DB 27 CAAAAAAC 34

```

RESULT 47
O4RVN9_TETNG PRELIMINARY; PRT; 691 AA.
ID O4RVN9_TETNG PRELIMINARY; PRT; 691 AA.
AC O4RVN9_TETNG PRELIMINARY; PRT; 691 AA.
DT 13-SEP-2005 (TEMBLrel. 31, Created)
DT 13-SEP-2005 (TEMBLrel. 31, Last sequence update)
DE Chromosome 10 SCAP15009, whole genome shotgun sequence.
GN ORFname=GSTENG00031390001;
OS Tetradodon nigroviridis (green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
NCBI_TaxId=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Saitanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,
RA Crnaud C., Duprat S., Brotilier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Lander E., Schachter V., Queller F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.,
RA "Genome duplication in the teleost fish tetradodon nigroviridis reveals
RT the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAE01015009; CACG09923.1; -; Genomic DNA.
SQ SEQUENCE 691 AA; 78163 MW; 440CAE5E405255CD CRC64;

Query Match 60.0%; Score 18; DB 2; Length 691;
Best Local Similarity 25.0%; Pred. No. 5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 288 CSSSSSSC 295

RESULT 48
O7S136_NEUCR PRELIMINARY; PRT; 705 AA.
ID O7S136_NEUCR PRELIMINARY; PRT; 705 AA.
AC O7S136_NEUCR PRELIMINARY; PRT; 705 AA.
DT 01-MAR-2004 (TEMBLrel. 26, Created)
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU09990.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
NCBI_TaxId=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Federsen D., Nelson M., Washburne M.,

```

```

RA Selltremlkoff C.P., Kinsey J.A., Braut E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Wewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysseil M., Mauceli E., Bieleke C., Rudd S., Fishman D.,
RA Kryzofova S., Raamussen C., Metzberg R.L., Perkins D.D., Kroen S.,
RA Cognoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmari S.A.,
RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannheim G., Ebbole D.J., Preitig M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
RL Nature 0:0-0(2003).
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000465; EAA29063.1; -; Genomic DNA.
SQ SEQUENCE 705 AA; 75444 MW; D7AB40BF0243E134 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 705;
Best Local Similarity 25.0%; Pred. No. 5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 367 CTSTATAC 374

RESULT 49
O5RFW2_BRARE PRELIMINARY; PRT; 711 AA.
ID O5RFW2_BRARE PRELIMINARY; PRT; 711 AA.
AC O5RFW2_BRARE PRELIMINARY; PRT; 711 AA.
DT 01-FEB-2005 (TEMBLrel. 29, Created)
DT 01-FEB-2005 (TEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TEMBLrel. 29, Last annotation update)
DE Novel protein (Zgc:77446).
GN Name=CH211-200P13.3; ORFNames=CH211-200P13.3-001, zgc:77446;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
NCBI_TaxId=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Donaldson S.,
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CHS36610; CA11747.1; -; Genomic DNA.
DR ZFIN; ZDB-GENE-040426-933; zgc:77446.
DR InterPro; IPR003892; CUF.
DR InterPro; IPR001876; Znf_RangDP.
DR Pfam; PF02845; CUF; 1.
DR SMART; SM00546; CUF; 1.
DR SMART; SM00547; Znf_RBZ; 1.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS0199; ZF_RANBP2_2; 1.
SQ SEQUENCE 711 AA; 76880 MW; E26723DFF30859356 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 711;
Best Local Similarity 25.0%; Pred. No. 5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 330 CESTSSC 337

RESULT 50
O5AP84_CANAL PRELIMINARY; PRT; 722 AA.
ID O5AP84_CANAL PRELIMINARY; PRT; 722 AA.
AC O5AP84_CANAL PRELIMINARY; PRT; 722 AA.
DT 10-MAY-2005 (TEMBLrel. 30, Created)
DT 10-MAY-2005 (TEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TEMBLrel. 30, Last annotation update)
DE Hypothetical YFM family protein 5.

```

GN Name=YFW5; ORFNames=CaO19.12344;
OS Candida albicans SC5314.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=237561;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
RX PubMed=15123810; DOI=10.1073/pnas.0401648101;
RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,
RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
RA Davis R.W., Scherer S.;
RT "The diploid genome sequence of Candida albicans."
RL Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
RA Dungan J., Kuo A., Newport G., Ian C.-Y., Iijima C., Adegbola O.,
RA Roberts J., Persson K., Donnelly S., Favoreto S., Tung K.-W.,
RA Jones T., Scherer S., Agabian N.;
RT "Annotation of the Genome of Candida albicans."
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAC001000002; EAL04673.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 722 AA; 81770 MW; 16E7C88A35A416AE CRC64;

Query Match 60.0%; Score 18; DB 2; Length 722;
Best local Similarity 25.0%; Pred. No. 5.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
DB 636 CTSATTSC 643

Search completed: January 4, 2006, 16:08:27
Job time : 124.426 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 : Search time 102.652 Seconds
(without alignments)
64.204 Million cell updates/sec

Title: US-09-932-322-3
Perfect score: 31
Sequence: 1 XXXXXXXXXXXXX 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21.*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	58.1	59	4	ABBA1147 Peptide #
2	18	58.1	59	4	ABBA1147 Peptide #
3	18	58.1	59	4	ABBA1147 Peptide #
4	18	58.1	59	4	ABBA1147 Peptide #
5	18	58.1	59	4	ABBA1147 Peptide #
6	18	58.1	59	4	ABBA1147 Peptide #
7	18	58.1	59	4	ABBA1147 Peptide #
8	18	58.1	59	4	ABBA1147 Peptide #
9	18	58.1	59	4	ABBA1147 Peptide #
10	18	58.1	59	4	ABBA1147 Peptide #
11	18	58.1	59	4	ABBA1147 Peptide #
12	18	58.1	59	4	ABBA1147 Peptide #
13	18	58.1	59	4	ABBA1147 Peptide #
14	18	58.1	59	4	ABBA1147 Peptide #
15	18	58.1	59	4	ABBA1147 Peptide #
16	18	58.1	59	4	ABBA1147 Peptide #
17	18	58.1	59	4	ABBA1147 Peptide #
18	18	58.1	59	4	ABBA1147 Peptide #
19	18	58.1	59	4	ABBA1147 Peptide #
20	18	58.1	59	4	ABBA1147 Peptide #
21	18	58.1	59	4	ABBA1147 Peptide #
22	18	58.1	59	4	ABBA1147 Peptide #
23	18	58.1	59	4	ABBA1147 Peptide #
24	18	58.1	59	4	ABBA1147 Peptide #

25	18	58.1	169	7	ABO83337 Pseudomon
26	18	58.1	174	8	ADP30801 Human sec
27	18	58.1	193	7	ABO73977 Pseudomon
28	18	58.1	198	8	ADP30493 Human sec
29	18	58.1	198	8	ADP30492 Human sec
30	18	58.1	198	8	ADP30477 Human sec
31	18	58.1	198	8	ADP30481 Human sec
32	18	58.1	198	8	ADP30491 Human sec
33	18	58.1	228	8	ADP31281 Human sec
34	18	58.1	229	7	ABO80501 Pseudomon
35	18	58.1	233	7	ABO77555 Pseudomon
36	18	58.1	234	8	ADP31468 Human sec
37	18	58.1	242	8	ADP31616 Human sec
38	18	58.1	246	5	AU75930 Androgen
39	18	58.1	246	7	ABO74848 Pseudomon
40	18	58.1	249	8	ADP30754 Human sec
41	18	58.1	254	8	ADP31336 Human sec
42	18	58.1	279	8	ADP30846 Human sec
43	18	58.1	279	8	ADP31523 Human sec
44	18	58.1	281	4	ABBA8050 Signal tr
45	18	58.1	281	4	ADBA3729 Corn DREB
46	18	58.1	286	7	ABO79000 Pseudomon
47	18	58.1	297	8	ADP31192 Human sec
48	18	58.1	312	8	ADP31454 Human sec
49	18	58.1	317	8	ADP31615 Human sec
50	18	58.1	339	8	ADP30519 Human sec
51	18	58.1	346	8	ADP31355 Human sec
52	18	58.1	348	8	ADP31441 Human sec
53	18	58.1	348	8	ADP30985 Human sec
54	18	58.1	348	8	ADP31676 Human sec
55	18	58.1	354	8	ADP30783 Human sec
56	18	58.1	354	8	ADP30779 Human sec
57	18	58.1	354	8	ADP30784 Human sec
58	18	58.1	354	8	ADP30778 Human sec
59	18	58.1	354	8	ADP31209 Human sec
60	18	58.1	360	8	ADP30659 Human sec
61	18	58.1	363	8	ADP30582 Human sec
62	18	58.1	369	8	ADP31492 Human sec
63	18	58.1	372	8	ADP31610 Human sec
64	18	58.1	372	8	ADP31221 Human sec
65	18	58.1	382	8	ADP31368 Human sec
66	18	58.1	396	8	ADP31378 Human sec
67	18	58.1	398	7	ABO71161 Pseudomon
68	18	58.1	405	8	ADP31534 Human sec
69	18	58.1	411	8	ADP30729 Human sec
70	18	58.1	426	8	ADP31495 Human sec
71	18	58.1	442	8	ADP31134 Human sec
72	18	58.1	444	8	ADP31640 Human sec
73	18	58.1	471	8	ADP31083 Human sec
74	18	58.1	474	8	ADP31083 Human sec
75	18	58.1	483	8	ADP30852 Human sec
76	18	58.1	483	8	ADP30853 Human sec
77	18	58.1	483	8	ADP31628 Human sec
78	18	58.1	489	8	ADP31380 Human sec
79	18	58.1	519	8	ADP31190 Human sec
80	18	58.1	525	8	ADP31227 Human sec
81	18	58.1	528	8	ADP30529 Human sec
82	18	58.1	528	8	ADP31525 Human sec
83	18	58.1	533	8	ADP31279 Human sec
84	18	58.1	543	8	ADP30872 Human sec
85	18	58.1	549	8	ADP30855 Human sec
86	18	58.1	552	8	ADP31524 Human sec
87	18	58.1	558	8	ADP31254 Human sec
88	18	58.1	558	8	ADP31255 Human sec
89	18	58.1	592	8	ADP30917 Human sec
90	18	58.1	603	8	ADP31645 Human sec
91	18	58.1	610	8	ADP31264 Human sec
92	18	58.1	615	8	ADP31169 Human sec
93	18	58.1	615	8	ADP31132 Human sec
94	18	58.1	615	8	ADP31361 Human sec
95	18	58.1	615	8	ADP30803 Human sec
96	18	58.1	615	8	ADP31360 Human sec
97	18	58.1	615	8	ADP31360 Human sec

244	18	58.1	1518	8	ADP31532	Human sec	317	18	58.1	2622	8	ADP31663	Adp31663 Human sec
245	18	58.1	1550	8	ADP30567	Human sec	318	18	58.1	2709	8	ADP31526	Adp31526 Human sec
246	18	58.1	1560	8	ADP31631	Human sec	319	18	58.1	2790	8	ADP31232	Adp31232 Human sec
247	18	58.1	1584	8	ADP31405	Human sec	320	18	58.1	2828	8	ADP30938	Adp30938 Human sec
248	18	58.1	1588	5	ABB09437	H. Influe	321	18	58.1	2833	8	ADP31299	Adp31299 Human sec
249	18	58.1	1614	8	ADP31529	Human sec	322	18	58.1	2835	8	ADP30572	Adp30572 Human sec
250	18	58.1	1617	8	ADP30660	Human sec	323	18	58.1	2976	9	AEA36049	Aea36049 Maize Sta
251	18	58.1	1647	8	ADP31052	Human sec	324	18	58.1	2980	9	AEA36049	Aea36049 Maize Sta
252	18	58.1	1662	8	ADP31419	Human sec	325	18	58.1	3036	8	ADP31595	Adp31595 Human sec
253	18	58.1	1662	8	ADP31513	Human sec	326	18	58.1	3144	8	ADP31544	Adp31544 Human sec
254	18	58.1	1665	8	ADP31187	Human sec	327	18	58.1	3201	8	ADP31545	Adp31545 Human sec
255	18	58.1	1719	8	ADP31137	Human sec	328	18	58.1	3339	8	ADP31219	Adp31219 Human sec
256	18	58.1	1725	8	ADP30654	Human sec	329	18	58.1	3350	8	ADP31148	Adp31148 Human sec
257	18	58.1	1743	6	ABU88255	Novel hum	330	18	58.1	3411	8	ADP30667	Adp30667 Human sec
258	18	58.1	1743	6	ABU90134	Novel hum	331	18	58.1	3447	8	ADP31112	Adp31112 Human sec
259	18	58.1	1743	6	ABU96436	Novel hum	332	18	58.1	3465	8	ADP31234	Adp31234 Human sec
260	18	58.1	1743	6	ABU99045	Novel hum	333	18	58.1	3579	8	ADP31028	Adp31028 Human sec
261	18	58.1	1743	6	ABU98260	Novel hum	334	18	58.1	4440	6	ABU88256	Abu88256 Novel hum
262	18	58.1	1743	6	ABU91966	Novel hum	335	18	58.1	4440	6	ABU90135	Abu90135 Novel hum
263	18	58.1	1743	6	ABU85270	Novel hum	336	18	58.1	4440	6	ABU96437	Abu96437 Novel hum
264	18	58.1	1743	6	ABO00409	Novel hum	337	18	58.1	4440	6	ABU99046	Abu99046 Novel hum
265	18	58.1	1743	6	ABU88960	Novel hum	338	18	58.1	4440	6	ABU98261	Abu98261 Novel hum
266	18	58.1	1743	6	ABO06456	Novel hum	339	18	58.1	4440	6	ABU91967	Abu91967 Novel hum
267	18	58.1	1743	6	ABU95516	Novel hum	340	18	58.1	4440	6	ABU85271	Abu85271 Novel hum
268	18	58.1	1743	6	ABU95206	Novel hum	341	18	58.1	4440	6	ABO00410	AbO00410 Novel hum
269	18	58.1	1743	6	ABU90754	Novel hum	342	18	58.1	4440	6	ABU88961	Abu88961 Novel hum
270	18	58.1	1743	6	ABU93916	Novel hum	343	18	58.1	4440	6	ABO06457	AbO06457 Novel hum
271	18	58.1	1743	6	ABU86190	Novel hum	344	18	58.1	4440	6	ABU95517	Abu95517 Novel hum
272	18	58.1	1743	6	ABU82045	Novel hum	345	18	58.1	4440	6	ABU95207	Abu95207 Novel hum
273	18	58.1	1743	6	ABU07906	Novel hum	346	18	58.1	4440	6	ABU90755	Abu90755 Novel hum
274	18	58.1	1743	6	ABU94226	Novel hum	347	18	58.1	4440	6	ABU93917	Abu93917 Novel hum
275	18	58.1	1743	6	ABO00099	Novel hum	348	18	58.1	4440	6	ABU86191	Abu86191 Novel hum
276	18	58.1	1743	6	ABU91351	Novel hum	349	18	58.1	4440	6	ABU82046	Abu82046 Novel hum
277	18	58.1	1743	6	ABU91351	Novel hum	350	18	58.1	4440	6	ABU07907	Abu07907 Novel hum
278	18	58.1	1743	6	ABU90444	Novel hum	351	18	58.1	4440	6	ABU94227	Abu94227 Novel hum
279	18	58.1	1743	6	ABU97035	Novel hum	352	18	58.1	4440	6	ABO00100	AbO00100 Novel hum
280	18	58.1	1743	6	ABO05231	Novel hum	353	18	58.1	4440	6	ABU87111	Abu87111 Novel hum
281	18	58.1	1746	8	ADP30992	Human sec	354	18	58.1	4440	6	ABU91352	Abu91352 Novel hum
282	18	58.1	1776	9	ADY62715	Human alp	355	18	58.1	4440	6	ABU90445	Abu90445 Novel hum
283	18	58.1	1782	8	ADP31391	Human sec	356	18	58.1	4440	6	ABU97036	Abu97036 Novel hum
284	18	58.1	1782	8	ADP31270	Human sec	357	18	58.1	4440	6	ABO05232	AbO05232 Novel hum
285	18	58.1	1789	8	ADP30962	Human sec	358	18	58.1	4752	8	ADP30585	Adp30585 Human sec
286	18	58.1	1793	4	ABB60964	Drosophi1	359	18	58.1	4752	8	ADP30651	Adp30651 Human sec
287	18	58.1	1815	8	ADP31601	Human sec	360	18	58.1	5304	8	ADP30706	Adp30706 Human sec
288	18	58.1	1827	8	ADP31170	Human sec	361	18	58.1	5397	8	ADP31068	Adp31068 Human sec
289	18	58.1	1833	8	ADP30642	Human sec	362	18	58.1	5514	8	ADP31186	Adp31186 Human sec
290	18	58.1	1933	8	ADP30889	Human sec	363	18	58.1	5514	8	ADP31591	Adp31591 Human sec
291	18	58.1	1933	8	ADP30902	Human sec	364	18	58.1	6465	8	ADP30705	Adp30705 Human sec
292	18	58.1	1956	8	ADP31662	Human sec	365	18	58.1	7285	6	ABJ38280	AbJ38280 pMMG21 -RA
293	18	58.1	2001	8	ADP31644	Human sec	366	18	58.1	7339	6	AAO16358	AaO16358 Human t1a
294	18	58.1	2020	8	ADP31056	Human sec	367	18	58.1	8976	8	ADP31425	Adp31425 Human sec
295	18	58.1	2058	8	ADP31630	Human sec	368	18	58.1	9195	8	ADP31494	Adp31494 Human sec
296	18	58.1	2065	4	ABB63705	Drosophi1	369	18	58.1	10944	8	ADP31311	Adp31311 Human sec
297	18	58.1	2088	8	ADP31178	Human sec	370	18	58.1	11328	4	ABBS6666	Abbs6666 Human SNP
298	18	58.1	2091	8	ADP31088	Human sec	371	17	54.8	13	4	ABBS6666	Abbs6666 Human SNP
299	18	58.1	2127	8	ADP31327	Human sec	372	17	54.8	23	7	AAE12409	Aae12409 Albunin f
300	18	58.1	2148	8	ADP30974	Human sec	373	17	54.8	23	7	ADP068071	Adp068071 Human the
301	18	58.1	2167	8	ADP30882	Human sec	374	17	54.8	23	9	AEA39569	Aea39569 Nematode
302	18	58.1	2304	8	ADP31252	Human sec	375	17	54.8	23	9	AEA39570	Aea39570 Nematode
303	18	58.1	2347	8	ADP31394	Human sec	376	17	54.8	23	9	AEA39571	Aea39571 Nematode
304	18	58.1	2349	8	ADP30959	Human sec	377	17	54.8	24	7	ADD95109	Add95109 PCR prime
305	18	58.1	2391	8	ADP31366	Human sec	378	17	54.8	24	8	ADU59677	AdU59677 GBP-4 c10
306	18	58.1	2418	8	ADP31105	Human sec	379	17	54.8	25	2	AAU21389	Aau21389 Human HUP
307	18	58.1	2454	8	ADP30469	Human sec	380	17	54.8	28	4	AAU33045	Aau33045 Peptide #
308	18	58.1	2484	8	ADP66690	Human sec	381	17	54.8	28	4	AAU72815	Aau72815 Human bon
309	18	58.1	2508	6	ADA15721	C. elegan	382	17	54.8	28	4	AAU60195	Aau60195 Human b1a
310	18	58.1	2535	8	ADP31146	Human sec	383	17	54.8	28	4	ABG54519	Abg54519 Human l1v
311	18	58.1	2542	8	ADP31594	Human sec	384	17	54.8	28	5	ABG42643	Abg42643 Human pep
312	18	58.1	2544	6	ADA15717	C. elegan	385	17	54.8	29	8	ADP30951	Adp30951 Human sec
313	18	58.1	2547	8	ADP31665	Human sec	386	17	54.8	31	3	AAU55811	Aau55811 Arabidops
314	18	58.1	2601	6	ADA15723	C. elegan	387	17	54.8	33	9	AEA34230	Aea34230 Opium pop
315	18	58.1	2616	8	ADP31253	Human sec	388	17	54.8	35	3	AAU99909	Aau99909 Peptide e
316	18	58.1	2616	9	AEA9675	N. mening	389	17	54.8	45	6	ABR98384	AbR98384 Tumour. ce


```
974 17 54.8 1173 8 ADP31155 Adp31155 Human sec
975 17 54.8 1179 8 ADP30486 Adp30486 Human sec
976 17 54.8 1179 8 ADP30577 Adp30577 Human sec
977 17 54.8 1183 8 ADP30536 Adp30536 Human sec
978 17 54.8 1192 8 ADP31180 Adp31180 Human sec
979 17 54.8 1194 8 ADP30682 Adp30682 Human sec
980 17 54.8 1197 8 ADP31034 Adp31034 Human sec
981 17 54.8 1197 8 ADP31342 Adp31342 Human sec
982 17 54.8 1200 9 ADZ76047 Adz76047 Mouse pre
983 17 54.8 1222 8 ADP30501 Adp30501 Human sec
984 17 54.8 1227 8 ADP31602 Adp31602 Human sec
985 17 54.8 1239 8 ADP31297 Adp31297 Human sec
986 17 54.8 1248 8 ADP31346 Adp31346 Human sec
987 17 54.8 1252 8 ADP30678 Adp30678 Human sec
988 17 54.8 1269 8 ADP31382 Adp31382 Human sec
989 17 54.8 1269 8 ADP31500 Adp31500 Human sec
990 17 54.8 1269 8 ADP31381 Adp31381 Human sec
991 17 54.8 1269 8 ADU23456 Adu23456 Sulfolobu
992 17 54.8 1282 8 ADP31328 Adp31328 Human sec
993 17 54.8 1289 8 ADP30675 Adp30675 Human sec
994 17 54.8 1300 6 ABU88254 Abu88254 Novel hum
995 17 54.8 1300 6 ABU90133 Abu90133 Novel hum
996 17 54.8 1300 6 ABU96435 Abu96435 Novel hum
997 17 54.8 1300 6 ABU99044 Abu99044 Novel hum
998 17 54.8 1300 6 ABU98259 Abu98259 Novel hum
999 17 54.8 1300 6 ABU91965 Abu91965 Novel hum
1000 17 54.8 1300 6 ABU85269 Abu85269 Novel hum
```

ALIGNMENTS

```
RESULT 1
ID ABB41147 standard; peptide; 59 AA.
XX ABB41147;
AC ABB41147;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #8653 encoded by human foetal liver single exon probe.
XX
KM Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human foetal liver.
XX
PS Claim 27; SEQ ID NO 31782; 639pp + Sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
XX human gene expression in a sample derived from human foetal liver. The
```

```
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human foetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at fcp.wipo.int/pub/published\_pct\_sequences
XX
SQ Sequence 59 AA;
```

```
Query Match 58.1%; Score 18; DB 4; Length 59;
Best Local Similarity 22.2%; Pred. No. 2.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
Qy 4 CXXXXXXC 12
Db 13 CXXSSTSTC 21
```

```
RESULT 2
ID AAM34923 standard; protein; 59 AA.
XX AAM34923;
AC AAM34923;
```

```
DT 17-OCT-2001 (first entry)
XX
DE Peptide #8960 encoded by probe for measuring placental gene expression.
XX
KM Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 35192; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SNP:
XX see AAI31315-AA157546). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders
XX
```

```
SQ Sequence 59 AA;
```

```
Query Match 58.1%; Score 18; DB 4; Length 59;
Best Local Similarity 22.2%; Pred. No. 2.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
Qy 4 CXXXXXXC 12
```

Db 13 CSSSSTSTC 21

RESULT 3
ID ABB25185 standard; protein; 59 AA.
AC ABB25185;
XX
XX 23-JAN-2002 (first entry)
XX
XX Protein #7184 encoded by probe for measuring heart cell gene expression.
DE
XX
XX Human; gene expression; heart; microarray; vascular system;
KM cardiovascular disease; hypertension; cardiac arrhythmia;
KM congenital heart disease.
XX
XX Homo sapiens.
OS
XX WO200157274-A2.
XX
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US000666.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488990/53.
DR
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
XX Claim 15; SEQ ID NO 26955; 530pp; English.
PS
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ffp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 59 AA;
SQ

Query Match 58.1%; Score 18; DB 4; Length 59;
Best Local Similarity 22.2%; Pred. No. 2.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CXXXXXXC 12
|
Db 13 CSSSSTSTC 21

RESULT 4
ID AAM74807 standard; protein; 59 AA.
XX
XX AAM74807;
AC

XX
DT 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 35113.
DE
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KM microarray; cancer; leukemia; lymphoma; myeloma.
XX
XX Homo sapiens.
OS
XX WO200157276-A2.
XX
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US000666.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488990/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 35113; 658pp + Sequence Listing; English.
PS
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
XX Sequence 59 AA;
SQ

Query Match 58.1%; Score 18; DB 4; Length 59;
Best Local Similarity 22.2%; Pred. No. 2.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CXXXXXXC 12
|
Db 13 CSSSSTSTC 21

RESULT 5
ID AAM62003 standard; protein; 59 AA.
XX
XX AAM62003;
AC
XX 05-NOV-2001 (first entry)
DT
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 34108.
DE
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
KM Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
XX Homo sapiens.
OS
XX WO200157275-A2.
XX
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US000667.
XX
XX

```

XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00613366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
XX Example 4; SEQ ID NO 34108; 650bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system,
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention
XX
XX Sequence 59 AA;
XX

```

```

Query Match          58.1%; Score 18; DB 4; Length 59;
Best Local Similarity 22.2%; Pred. No. 2.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

Oy 4 CXXXXXXC 12
Db 13 CXXXXXXC 21

```

```

RESULT 6
ABG56589
ID ABG56589 standard; peptide; 59 AA.
XX
XX ABG56589;
XX
XX 25-FEB-2003 (first entry)
XX
XX Human liver peptide, SEQ ID No 35237.
XX
XX Human, liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
XX hypercholesterolaemia; coronary heart disease.
XX
XX Homo sapiens.
XX
XX WO200157273-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000664.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00613366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-48898/53.

```

```

XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
XX Claim 27; SEQ ID NO 35237; 658bp; English.
XX

```

```

XX The invention relates to a single exon nucleic acid probe (SENP) (i) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridizes at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (ii) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABG47348-ABG5930 represent human
XX liver single exon encoded peptides of the invention. Note: The sequence
XX information for this patent does not appear in the printed specification
XX but was obtained in electronic format directly from WIPO at
XX ftp://ipub/pub/published_pct_sequences
XX

```

```

XX Sequence 59 AA;
XX

```

```

Query Match          58.1%; Score 18; DB 4; Length 59;
Best Local Similarity 22.2%; Pred. No. 2.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

Oy 4 CXXXXXXC 12
Db 13 CXXXXXXC 21

```

```

RESULT 7
ABG44598
ID ABG44598 standard; peptide; 59 AA.
XX
XX ABG44598;
XX
XX 19-AUG-2002 (first entry)
XX
XX Human peptide encoded by genome-derived single exon probe SEQ ID 34263.
XX
XX Human, single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.
XX
XX Homo sapiens.
XX
XX WO200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US000665.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00613366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.

```


XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 27; SEQ ID NO 34263; 634bp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridize at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarray having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Kargener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a peptide/protein encoded by a single exon probe of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pcr_sequences
XX
SQ Sequence 59 AA;
XX
XX Query Match 58.1%; Score 18; DB 5; Length 59;
XX Best Local Similarity 22.2%; Pred. No. 2.8e+02;
XX Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
QY 4 CXXXXXXC 12
XX |
DB 13 CXXXXXXC 21

RESULT 8
AAB21190
ID AAB21190 standard; protein; 108 AA.
XX
AC AAB21190;
XX
XX 08-JAN-2001 (first entry)
XX
XX Exo14 partial protein.
XX
XX Exocytosis pathway protein; Exo protein; antiparkinsonian; antidiabetic;
XX antiallergic; antiaslatic; neuroprotective; anticonvulsant;
XX vulnerable; asthma; inflammation; allergy; Chediak-Higashi syndrome; CHS;
XX Alzheimer's disease; Parkinson's disease; Huntington's disease; diabetes;
XX digestion disorder; wound healing disorder; gene therapy.
XX
OS Mus sp.

XX
XX WO200043419-A2.
XX
XX 27-JUL-2000.
XX
XX 20-JAN-2000; 2000WO-US001431.
XX
XX 20-JAN-1999; 99US-0116534P.
XX 26-JAN-1999; 99US-0117274P.
XX 26-JAN-1999; 99US-0117308P.
XX 26-JAN-1999; 99US-0117309P.
XX 26-JAN-1999; 99US-0117312P.
XX 01-FEB-1999; 99US-0118177P.
XX 01-FEB-1999; 99US-0118178P.
XX 01-FEB-1999; 99US-0118179P.
XX 09-FEB-1999; 99US-0119286P.
XX 11-FEB-1999; 99US-0119759P.
XX 11-FEB-1999; 99US-0119988P.
XX
XX (RIGEL-) RIGEL PHARM INC.
XX
XX Luo Y;
XX
XX WPI; 2000-482908/42.
XX
XX N-PSDB; AAA89569.
XX
XX New nucleic acids encoding Exo proteins which are useful in the
XX diagnosis, treatment or prevention of exocytosis-mediated disorders such
XX as asthma, inflammation and allergies.
XX
XX Disclosure; Page 149-150; 305pp; English.
XX
XX The present sequence is a polypeptide which is associated with the
XX exocytosis pathway. cDNA molecules encoding proteins involved in
XX exocytosis have been isolated by yeast one-hybrid and two-hybrid
XX screening. Novel proteins, termed Exo proteins, have been identified that
XX interact with known exocytosis-associated proteins such as GS27, alpha
XX snap, unc18-1, vamps, snap-23, and the rab family of proteins. Exo
XX proteins and their agonists and antagonists are useful in the diagnosis,
XX treatment or prevention of exocytosis-mediated disorders such as asthma,
XX inflammation, allergies, Chediak-Higashi Syndrome (CHS), Alzheimer's
XX disease, Parkinson's disease, Huntington's disease, diabetes, digestion
XX disorders and wound healing disorders. The nucleic acids, antagonists or
XX agonists of Exo proteins are useful in gene therapy. The nucleic acids
XX are also useful for generating transgenic or knock-out animals which can
XX be used in the development and screening of therapeutically useful
XX reagents
XX
XX Sequence 108 AA;
XX
XX Query Match 58.1%; Score 18; DB 3; Length 108;
XX Best Local Similarity 22.2%; Pred. No. 3.2e+02;
XX Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
QY 4 CXXXXXXC 12
XX |
DB 33 CXXXXXXC 41

RESULT 9
ABO75321
ID ABO75321 standard; protein; 134 AA.
XX
XX ABO75321;
XX
XX 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polypeptide #7496.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
XX
XX

PN US6551795-B1.
XX
XX 22-APR-2003.
PD
XX
PF 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX
XX 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX N-PSDB; ABD08892.
DR
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
PS
XX Disclosure; SEQ ID NO 24067; 455bp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences AB067826-
XX AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX Sequence 134 AA;
SQ

Query Match 58.1%; Score 18; DB 7; Length 134;
Best Local Similarity 22.2%; Pred. No. 3.4e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CXXXXXXXC 12
DB 19 CCASTTSC 27

RESULT 10
ADE72390
ID ADE72390 standard; protein; 139 AA.
XX
XX ADE72390;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX Human endometrial specific protein, SEQ ID NO 430.
DE
XX
XX cytotatic; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer.
XX
XX Homo sapiens.
OS
XX
XX WO2003060081-A2.
XX
XX 24-JUL-2003.
PD
XX
XX 23-DEC-2002; 2002WO-US041612.
XX
XX 21-DEC-2001; 2001US-0342756P.
XX

PA (DIAD-) DIADEXUS INC.
XX
XX Sun Y, Liu C;
XX
XX WPI; 2003-577666/54.
XX
XX Nucleic acid molecules and polypeptides useful for diagnosing and
XX treating endometrial cancer and non-cancerous disease states in
XX endometrial.
XX
XX Claim 12; SEQ ID NO 430; 824bp; English.
XX
XX The invention comprises the amino acid and DNA sequences of human
XX endometrial specific genes and proteins. The DNA and protein sequences of
XX the invention are useful for diagnosing, imaging and treating a patient
XX with endometrial cancer. The present amino acid sequence represents a
XX human endometrial specific protein of the invention.
XX
XX Sequence 139 AA;
SQ

Query Match 58.1%; Score 18; DB 7; Length 139;
Best Local Similarity 22.2%; Pred. No. 3.4e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CXXXXXXXC 12
DB 60 CSTSSAASC 68

RESULT 11
ADE72391
ID ADE72391 standard; protein; 139 AA.
XX
XX ADE72391;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX Human endometrial specific protein, SEQ ID NO 431.
DE
XX
XX cytotatic; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer.
XX
XX Homo sapiens.
OS
XX
XX WO2003060081-A2.
XX
XX 24-JUL-2003.
PD
XX
XX 23-DEC-2002; 2002WO-US041612.
XX
XX 21-DEC-2001; 2001US-0342756P.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Sun Y, Liu C;
XX
XX WPI; 2003-577666/54.
XX
XX Nucleic acid molecules and polypeptides useful for diagnosing and
XX treating endometrial cancer and non-cancerous disease states in
XX endometrial.
XX
XX Claim 12; SEQ ID NO 431; 824bp; English.
XX
XX The invention comprises the amino acid and DNA sequences of human
XX endometrial specific genes and proteins. The DNA and protein sequences of
XX the invention are useful for diagnosing, imaging and treating a patient
XX with endometrial cancer. The present amino acid sequence represents a
XX human endometrial specific protein of the invention.
XX
XX Sequence 139 AA;
SQ

Query Match 58.1%; Score 18; DB 7; Length 139;

Best Local Similarity 22.2%; Pred. No. 3.4e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 12
Db 60 CSTRSASAC 68

RESULT 12

AAM93094
ID AAM93094 standard; protein; 140 AA.

AC AAM93094;

DT 06-NOV-2001 (first entry)

DE Human digestive system antigen SEQ ID NO: 2443.

KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum.

OS Homo sapiens.

PN WO200155314-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US001324.

PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234224P.

PR 25-SEP-2000; 2000US-0234974P.

PR 25-SEP-2000; 2000US-0234977P.

PR 26-SEP-2000; 2000US-0234988P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239335P.

PR 13-OCT-2000; 2000US-0239337P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.

PR 17-NOV-2000; 2000US-0249210P.

PR 17-NOV-2000; 2000US-0249211P.

PR 17-NOV-2000; 2000US-0249212P.

PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.

PR 17-NOV-2000; 2000US-0249215P.

PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.

PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-502630/55.
DR N-PSDB; AAK88867.
XX
XX Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognostizing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX
XX Claim 11; SEQ ID NO 2443; 986bp; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a digestive system antigen of
CC the invention
XX
XX SQ Sequence 140 AA;
Query Match 58.1%; Score 18; DB 4; Length 140;
Best Local Similarity 22.2%; Pred. No. 3.4e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CXXXXXXC 12
DB 104 CTATSATAC 112
RESULT 13
ID AAU20102 standard; protein; 140 AA.
XX
XX AAU20102;
DT 06-DEC-2001 (first entry)
XX
DE Human liver associated polypeptide #133.
XX
XX Liver associated protein; human; mouse; rabbit; goat; horse; cat; dog;
XX chicken; sheep; immunosuppressive; antihairtic; vasotropic;
XX antihemetic; antiproliferative; cytostatic; cardiant; neuroprotective;
XX cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
XX ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;
XX hyperproliferative disorder; breast; liver; cardiovascular disorder;
XX cerebrovascular disorder; nervous system disorder; bacterial infection;
XX fungal infection; viral infection; ocular disorder; endocrine disorder;
XX gastrointestinal disorder; renal disorder; respiratory disorder;
XX wound healing; skin aging; organ transplantation; tissue regeneration;

KW anti-infertility.
XX
XX Homo sapiens.
OS
XX
XX WO200155355-A1.
XX
XX 02-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US001351.
PF
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0196123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-021647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225470P.
PR 14-AUG-2000; 2000US-022547P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-02331968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.

```

PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236337P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239937P.
PR 13-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241888P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251888P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251900P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-457728/49.
XX
XX N-PSDB; AAS31813.
XX
XX Isolated nucleic acid molecule encoding a human liver related protein is
PT used in preventing, treating or ameliorating disorders of the liver
PT particularly cancer of the liver.
XX
XX Claim 11; SEQ ID NO 289; 526pp; English.
XX
XX Sequences AAU19970-AAU20115 represent the liver associated polypeptides
CC of the invention. Liver associated polypeptides and their associated
CC polynucleotides are useful in the diagnosis, treatment and prevention of
CC various types of disorders in e.g. humans, mice, rabbits, goats, horses,
CC cats, dogs, chickens or sheep. A pathological condition can be determined
CC by detecting the presence or absence of a mutation in a liver associated
CC polynucleotide. The treatable disorders include autoimmune diseases such
CC as rheumatoid arthritis, hyperproliferative disorders such as neoplasms
CC of the breast or liver, cardiovascular disorders such as cardiac arrest,
CC cerebrovascular disorders such as cerebral ischaemia, nervous system
CC disorders such as Alzheimer's disease, infections caused by bacteria,
CC viruses and fungi, ocular disorders such as corneal infection, endocrine
CC disorders such as Crohn's disease, renal disorders such as
CC glomerulonephritis and respiratory disorders such as asthma and pleurisy.
CC The polypeptides can also be used to aid wound healing, to prevent skin
CC aging due to sunburn, to maintain organs before transplantation, to
CC regenerate tissues and in chemotaxis. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
Query Match 58.1%; Score 18; DB 4; Length 140;
Best Local Similarity 22.2%; Pred. No. 3.4e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CXXXXXXXC 12
Db 104 CTATSATAC 112
RESULT 14
ABP40963
ID ABP40963 standard; protein; 140 AA.
XX
AC ABP40963;
XX
XX 24-JUL-2002 (first entry)
XX
XX Human liver antigen HFLV70, SEQ ID NO:289.
DB
XX
XX Human; liver antigen; liver disorder; hepatic disorder; infection;
KW hepatitis; viral; parasitic; bacterial; fungal; inflammatory condition;
KW cirrhosis; granulomatous hepatitis; toxin damage; drug damage;
KW autoimmune disease; Wilson's disease; primary biliary cirrhosis;
KW neoplastic disorder; cancer; tumour; portal hypertension;
KW gastrointestinal disorder; hepatitis; drug screening; gene therapy;
KW chromosome mapping; forensic analysis; antibody preparation;
KW hepatotropic; cytostatic; antiinflammatory; vincristine; antibacterial;
KW fungicide; parasiticicide; antidote; immunosuppressive.
XX
XX Homo sapiens.
OS
XX
XX US2002042096-A1.
XX
XX 11-APR-2002.
XX
XX 17-JAN-2001; 2001US-00764887.

```

XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216680P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226686P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234423P.
PR 21-SEP-2000; 2000US-0234423P.
PR 25-SEP-2000; 2000US-0234977P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2002-381944/41.
DR N-PSDB; ABN90168.
XX
PT New nucleic acid encoding human liver antigens, useful for diagnosis,
PT treatment and prevention of e.g. hepatitis and hepatic cancer, also
XX related polypeptides and antibodies.
XX
PS Claim 11; SEQ ID NO 289; 181bp; English.
XX
CC The invention relates to 145 novel human liver antigens (ABP40831-
CC ABP40975) and to cDNAs encoding them (ABN90036-ABN90180), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human liver antigen
CC polynucleotides, antibodies against human liver antigens, and the use of

CC liver antigen polynucleotides and polypeptides in diagnosing, treating,
CC prognosing or preventing various disorders of the liver. Such conditions
CC include viral infections (e.g., cytomegalovirus, Epstein-Barr virus),
CC hepatitis A virus, hepatitis B virus and hepatitis C virus), parasitic
CC infections (e.g., Clonorchis sinensis, Echinococcus granulosus and
CC Entamoeba histolytica), and also bacterial and fungal infections. Other
CC disorders that may be treated include inflammatory conditions (e.g.,
CC cirrhosis and granulomatous hepatitis), damage caused by drugs or toxins,
CC autoimmune diseases (e.g., Wilson's disease, primary biliary cirrhosis),
CC neoplastic disorders (e.g., adenomas, haemangiomas and hepatocellular
CC carcinoma), portal hypertension, or gastrointestinal disorders (e.g.,
CC peptic ulcers, gastritis and peritoneal diseases). Liver antigen
CC polypeptides and polynucleotides may also be used in screening for
CC compounds which modulate liver antigen expression or activity. The
CC polynucleotides may further be used for gene therapy, chromosome mapping,
CC in the identification of individuals and in forensic analysis, and the
CC polypeptides may be used as molecular weight markers or to prepare
CC antibodies useful in disease diagnosis, drug targeting and phenotyping.
CC The present sequence represents a human liver antigen of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO at seqdata.uspto.gov/sequence/.
XX
SQ Sequence 140 AA;
Query Match 58.1%; Score 18; DB 5; Length 140;
Best Local Similarity 22.2%; Pred. No. 3.4e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 4 CXXXXXXC 12
Db 104 CTATATAC 112
RESULT 15
ADJ15081
ID ADJ15081 standard; protein; 140 AA.
AC ADJ15081;
XX 20-MAY-2004 (first entry)
DT
XX
XX Human liver-related protein - SEQ ID 289.
XX
XX liver; virucide; fungicide; antibacterial; antiparasitic; hepatotropic;
XX antiinflammatory; cyrostatic; litholytic; antineumatic; antiarthritic;
XX neuroprotective; antidiabetic; anticoagulant; thrombolytic;
XX antiarteriosclerotic; cardiant; haemostatic; antiarrhythmic;
XX ophthalmological; antiarteriosclerotic; vasotropic; osteopathic;
XX nootropic; antiparkinsonian; anticonvulsant; neuroleptic; vasotropic;
XX cyrostatic; gynaecological; viral; fungal; bacterial;
XX parasitic infection; cirrhosis; Wilson's disease;
XX gastrointestinal disorder; pancreatic; gallbladder; immune; blood;
XX hyperproliferative; cardiovascular; respiratory; musculoskeletal system;
XX neurological; endocrine; reproductive system; developmental; inherited;
XX human.
XX
XX Homo sapiens.
OS
XX
XX US2003077602-A1.
XX
XX PD 24-APR-2003.
XX
XX PF 14-FEB-2002; 2002US-00073961.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.

07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764897.

(HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX
PI
PI
DR
DR
XX
XX
PT
PT
XX
XX
PS
XX
CC
CC
CC
CC
CC
antibacterial, antiparasitic, hepatotropic, antiinflammatory, cytostatic,
Claim 11; SEQ ID NO 269; 181bp; English.
The invention relates to a novel isolated, liver related polypeptide. The
polypeptide of the invention demonstrates virucide, fungicide,
antibacterial, antiparasitic, hepatotropic, antiinflammatory, cytostatic,

CC litholytic, antirheumatic, antiarthritic, neuroprotective, antidiabetic,
CC anticoagulant, thrombolytic, antiarteriosclerotic, cardiant, haemostatic,
CC antiarrhythmic, ophthalmologic, antiarteriosclerotic, vasotropic,
CC osteopathic, nootropic, antiparkinsonian, anticonvulsant, neuroleptic,
CC vasotropic, cytostatic and gynaecological activities. The polypeptides
CC and polynucleotides of the invention may be useful for diagnosis
CC detection, treatment and/or prevention of disorders of the liver such as
CC viral, fungal, bacterial or parasitic infections, cirrhosis, Wilson's
CC disease, gastrointestinal disorders, pancreatic disorders, gallbladder
CC diseases, immune disorders, blood related disorders, hyperproliferative
CC disorders, cardiovascular disorders, respiratory disorders,
CC musculoskeletal system disorders, neurological diseases, endocrine
CC disorders, reproductive system disorders or developmental and inherited
CC disorders. The current sequence is that of the human liver-related
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was obtained electronically from the USPTO web-
CC site.

Query Match 58.1%; Score 18; DB 7; Length 140;
Best Local Similarity 22.2%; Pred. No. 3.4e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 4 CXXXXXXC 12
Db 104 CTRATSATAC 112

RESULT 16
ADP31453
ID ADP31453 standard; protein; 144 AA.

AC ADP31453;

DT 12-AUG-2004 (first entry)
XX Human secreted protein SEQ ID #2220.

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

PN WO2004035732-A2.

PD 29-APR-2004.

PF 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471306P.
PR 22-MAY-2003; 2003US-0472420P.
PR 09-JUN-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 14-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486446P.
PR 15-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486901P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 3451; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.

XX Sequence 144 AA;

Query Match 58.1%; Score 18; DB 8; Length 144;
Best Local Similarity 22.2%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485233P.
PR 08-JUL-2003; 2003US-0485244P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-048646P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493373P.
PR 08-AUG-2003; 2003US-0493377P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX William LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisnan L, Linnemann T;
PI Pierce K, Wang Y, Wong JG, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 3509; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases.
CC The present invention represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
SQ
SQ Sequence 153 AA;
XX
XX
Query Match 58.1%; Score 18; DB 8; Length 153;
Best Local Similarity 22.2%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CXXXXXXXXC 12
Db 79 CTATTTTTC 87
RESULT 20
ADN23734
ID ADN23734 standard; protein; 155 AA.
XX
XX ADN23734;
AC
XX 02-DEC-2004 (first entry)
DT XX
DE Bacterial polypeptide #6387.
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
XX Bacteria.
OS
XX
XX US2003233675-A1.
PN
XX
XX 18-DEC-2003.
PD
XX 20-FEB-2003; 2003US-00369493.
PF
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX DR NPI; 2004-061375/06.
XX PT New recombinant DNA construct comprising a promoter positioned to provide
XX PT expression of a polynucleotide encoding a polypeptide from a
XX PT microbial source, useful for producing plants with improved properties.
PS Claim 1; SEQ ID NO 6387; 122pp; English.
XX CC The invention relates to a recombinant DNA construct comprising a
XX CC promoter functional in a plant cell, where the promoter is positioned to
XX CC provide for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source. The invention also relates to a transformed plant
XX CC comprising the recombinant DNA construct and a method of producing a
XX CC transformed plant having an improved property. The plant is a crop plant
XX CC such as maize or soybean. The method of producing a transformed plant
XX CC having an improved property comprises transforming a plant with the
XX CC recombinant DNA construct and growing the transformed plant, where the
XX CC polynucleotide or polypeptide is useful for improving plant properties.
XX CC The recombinant DNA construct is useful for producing plants with
XX CC improved plant properties, e.g. improved cold, heat or drought tolerance,
XX CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX CC increased resistance to plant disease, better growth rate by modification
XX CC of the cell cycle pathway with plant growth regulators, increased rate of
XX CC homologous recombination, modified seed oil or protein yield and/or
XX CC content, improved yield by modification of carbohydrate, nitrogen or
XX CC phosphorus use and/or uptake, by modification of photosynthesis or by
XX CC providing improved plant growth and development under at least one stress
XX CC condition, improved lignin production or improved galactomannan
XX CC production. This sequence represents a bacterial polypeptide used in the
XX CC scope of the invention. Note: The sequence data for this patent did not
XX CC form part of the printed specification but was obtained in electronic
XX CC format from USPTO at seqdata.uspto.gov/sequence.html.
SQ Sequence 155 AA;
Query Match 58.1%; Score 18; DB 8; Length 155;
Best Local Similarity 22.2%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CXXXXXXXXC 12
DB 92 CTTASSSSC 100
RESULT 21
ABO83103
ID ABO83103 standard; protein; 156 AA.
XX AC ABO83103;
XX DT 29-UTL-2004 (first entry)
XX DE Pseudomonas aeruginosa polypeptide #15278.
XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US651795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX DR WPI; 2003-615309/58.

DR N-PSDB; ABD16674.
XX XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.
PS Disclosure; SEQ ID NO 31849; 455pp; English.
XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the
XX CC polynucleotides encoding them. The sequences are useful in diagnosis and
XX CC therapy of pathological conditions, as molecular targets for diagnostics,
XX CC prophylaxis and treatment of pathological conditions resulting from a
XX CC bacterial infection, for evaluating a compound, such as a polypeptide,
XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of
XX CC effective antibacterial targets, as targets for antibacterial drugs,
XX CC including anti-P. aeruginosa drugs, as templates for recombinant
XX CC production of P. aeruginosa-derived peptides or polypeptides, as target
XX CC components for diagnosis and/or treatment of P. aeruginosa-caused
XX CC infection, and in detection of P. aeruginosa sequences or other sequences
XX CC of Pseudomonas species using bioclip technology. Sequences ABO67826-
XX CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format from USPTO at
XX CC seqdata.uspto.gov/sequence.html.
SQ Sequence 156 AA;
Query Match 58.1%; Score 18; DB 7; Length 156;
Best Local Similarity 22.2%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CXXXXXXXXC 12
DB 46 CASAAATAC 54
RESULT 22
ADP30759
ID ADP30759 standard; protein; 156 AA.
XX AC ADP30759;
XX DT 12-AUG-2004 (first entry)
XX DE Human secreted protein SEQ ID #1526.
XX KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX KW cancer; inflammatory; immune; human secreted protein.
XX OS Homo sapiens.
XX PN WO2004035732-A2.
XX PD 29-APR-2004.
XX PF 28-AUG-2003; 2003WO-US026780.
XX PR 29-AUG-2002; 2002US-0406576P.
XX PR 29-AUG-2002; 2002US-0406579P.
XX PR 29-AUG-2002; 2002US-0406585P.
XX PR 29-AUG-2002; 2002US-0406588P.
XX PR 29-AUG-2002; 2002US-0406608P.
XX PR 29-AUG-2002; 2002US-0406611P.
XX PR 29-AUG-2002; 2002US-0406612P.
XX PR 29-AUG-2002; 2002US-0406616P.
XX PR 29-AUG-2002; 2002US-0406640P.
XX PR 29-AUG-2002; 2002US-0406642P.
XX PR 29-AUG-2002; 2002US-0406646P.
XX PR 29-AUG-2002; 2002US-0406653P.
XX PR 29-AUG-2002; 2002US-0406655P.
XX PR 29-AUG-2002; 2002US-0406666P.
XX PR 17-SEP-2002; 2002US-0410946P.
XX PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471366P.
PR 19-MAY-2003; 2003US-0471366P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX MPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2757; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present

CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 156 AA;
Query Match 58.1%; Score 18; DB 8; Length 156;
Best Local Similarity 22.2%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CXXXXXXXXC 12
DB 47 CTAATTTC 55
RESULT 23
ADP31286
ID ADP31286 standard; protein; 165 AA.
XX
AC ADP31286;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2053.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
PD 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
PF
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467206P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471366P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halembeck RF, Huang MM, Kochakota S, Haisnan L, Linnemann T,
PI Pierce X, Wang Y, Wong JGP, Wu G, Zhang H;
XX MPI; 2004-348438/32.

PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.

PS Claim 1; SEQ ID NO 3284; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.

XX Sequence 165 AA;

Query Match 58.1%; Score 18; DB 8; Length 165;
Best Local Similarity 22.2%; Pred. No. 3.6e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CXXXXXXXC 12
Db 107 CTTTAATC 115

RESULT 24
ADP31174
ID ADP31174 standard; protein; 165 AA.

XX ADP31174;
AC
XX
DT 12-AUG-2004 (first entry)
XX

DE Human secreted protein SEQ ID #1941.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW Cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

PN WO2004035732-A2.

XX 29-APR-2004.

XX 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406640P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406646P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

XX 17-SEP-2002; 2002US-0410946P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410957P.

XX 17-SEP-2002; 2002US-0410958P.

XX 17-SEP-2002; 2002US-0410959P.

XX 17-SEP-2002; 2002US-0410960P.

XX 17-SEP-2002; 2002US-0410961P.

XX 17-SEP-2002; 2002US-0410962P.

XX 17-SEP-2002; 2002US-0411019P.

XX 17-SEP-2002; 2002US-0411022P.

XX 17-SEP-2002; 2002US-0411023P.

XX 17-SEP-2002; 2002US-0411024P.

XX 17-SEP-2002; 2002US-0411032P.

XX 17-SEP-2002; 2002US-0411035P.

XX 17-SEP-2002; 2002US-0411037P.

XX 17-SEP-2002; 2002US-0411041P.

XX 17-SEP-2002; 2002US-0411045P.

XX 17-SEP-2002; 2002US-0411046P.

XX 17-SEP-2002; 2002US-0411048P.

XX 17-SEP-2002; 2002US-0411052P.

XX 17-SEP-2002; 2002US-0411055P.

XX 17-SEP-2002; 2002US-0411073P.

XX 17-SEP-2002; 2002US-0411073P.

XX 17-SEP-2002; 2002US-0411082P.

XX 17-SEP-2002; 2002US-0411101P.

XX 18-APR-2003; 2003US-0463700P.

XX 18-APR-2003; 2003US-0463708P.

XX 18-APR-2003; 2003US-0463716P.

XX 18-APR-2003; 2003US-0463732P.

XX 02-MAY-2003; 2003US-0467199P.

XX 02-MAY-2003; 2003US-0467201P.

XX 02-MAY-2003; 2003US-0467203P.

XX 02-MAY-2003; 2003US-0467230P.

XX 19-MAY-2003; 2003US-0471306P.

XX 19-MAY-2003; 2003US-0471366P.

XX 22-MAY-2003; 2003US-0472420P.

XX 09-JUN-2003; 2003US-0472430P.

XX 09-JUN-2003; 2003US-0476609P.

XX 08-JUL-2003; 2003US-0485218P.

XX 08-JUL-2003; 2003US-0485223P.

XX 08-JUL-2003; 2003US-0485224P.

XX 08-JUL-2003; 2003US-0485325P.

PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493373P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisnan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3172; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
XX Sequence 165 AA;
SQ

Query Match 58.1%; Score 18; DB 8; Length 165;
Best Local Similarity 22.2%; Pred. No. 3.6e+02; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 7;

QY 4 CXXXXXXXC 12
DB 107 CTTTAATC 115

RESULT 25
ABO83337
ID ABO83337 standard; protein; 169 AA.
XX
XX ABO83337;
AC
XX
XX 29-JUL-2004 (first entry)
DT
XX
XX Pseudomonas aeruginosa polypeptide #15512.
DE
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
KM
XX
XX Pseudomonas aeruginosa.
OS
XX
XX US651795-B1.
PN
XX
XX 22-APR-2003.
PD
XX
XX 18-FEB-1999; 99US-00252991.
PF
XX
XX 18-FEB-1998; 98US-0074788P.
PR
XX
XX 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI
XX
XX WPI; 2003-615309/58.
DR
XX
XX N-PSDB; ABD16908.
PR

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX
XX Disclosure; SEQ ID NO 32083; 455pp; English.
PS

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX

SQ Sequence 169 AA;

Query Match 58.1%; Score 18; DB 7; Length 169;
Best Local Similarity 22.2%; Pred. No. 3.6e+02; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 7;

QY 4 CXXXXXXXC 12
DB 84 CTTTSAAC 92

RESULT 26
ADP30801
ID ADP30801 standard; protein; 174 AA.
XX
XX ADP30801;
AC
XX
XX 12-AUG-2004 (first entry)
DT
XX
XX Human secreted protein SEQ ID #1568.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
OS
XX
XX WO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PF
XX
XX 29-AUG-2002; 2002US-0406576P.
PR
XX
XX 29-AUG-2002; 2002US-0406579P.
PR
XX
XX 29-AUG-2002; 2002US-0406585P.
PR
XX
XX 29-AUG-2002; 2002US-0406588P.
PR
XX
XX 29-AUG-2002; 2002US-0406608P.
PR
XX
XX 29-AUG-2002; 2002US-0406611P.
PR
XX
XX 29-AUG-2002; 2002US-0406612P.
PR
XX
XX 29-AUG-2002; 2002US-0406616P.
PR
XX
XX 29-AUG-2002; 2002US-0406640P.
PR
XX
XX 29-AUG-2002; 2002US-0406642P.
PR
XX
XX 29-AUG-2002; 2002US-0406646P.
PR
XX
XX 29-AUG-2002; 2002US-0406653P.
PR
XX
XX 29-AUG-2002; 2002US-0406655P.
PR
XX
XX 29-AUG-2002; 2002US-0406666P.
PR
XX
XX 17-SEP-2002; 2002US-0410946P.
PR
XX
XX 17-SEP-2002; 2002US-0410947P.
PR
XX
XX 17-SEP-2002; 2002US-0410948P.
PR
XX
XX 17-SEP-2002; 2002US-0410949P.
PR

Db	61	CTSTSSATC	69
RESULT_28			
ID	ADP30493		
XX	ADP30493	standard; protein; 198 AA.	
AC	ADP30493;		
DT	12-AUG-2004	(first entry)	
XX			
DE	Human secreted protein SEQ ID #1260.		
KW	Cycostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide; cancer; inflammatory; immune; human secreted protein.		
OS	Homo sapiens.		
XX			
PN	WO2004035732-A2.		
PD	29-APR-2004.		
PF	28-AUG-2003; 2003WO-US026780.		
XX			
PR	29-AUG-2002; 2002US-0406576P.		
PR	29-AUG-2002; 2002US-0406579P.		
PR	29-AUG-2002; 2002US-0406585P.		
PR	29-AUG-2002; 2002US-0406588P.		
PR	29-AUG-2002; 2002US-0406608P.		
PR	29-AUG-2002; 2002US-0406611P.		
PR	29-AUG-2002; 2002US-0406612P.		
PR	29-AUG-2002; 2002US-0406616P.		
PR	29-AUG-2002; 2002US-0406640P.		
PR	29-AUG-2002; 2002US-0406642P.		
PR	29-AUG-2002; 2002US-0406646P.		
PR	29-AUG-2002; 2002US-0406653P.		
PR	29-AUG-2002; 2002US-0406655P.		
PR	29-AUG-2002; 2002US-0406666P.		
PR	17-SEP-2002; 2002US-0410946P.		
PR	17-SEP-2002; 2002US-0410947P.		
PR	17-SEP-2002; 2002US-0410948P.		
PR	17-SEP-2002; 2002US-0410949P.		
PR	17-SEP-2002; 2002US-0410953P.		
PR	17-SEP-2002; 2002US-0410957P.		
PR	17-SEP-2002; 2002US-0410958P.		
PR	17-SEP-2002; 2002US-0410959P.		
PR	17-SEP-2002; 2002US-0410960P.		
PR	17-SEP-2002; 2002US-0410961P.		
PR	17-SEP-2002; 2002US-0410962P.		
PR	17-SEP-2002; 2002US-0411019P.		
PR	17-SEP-2002; 2002US-0411022P.		
PR	17-SEP-2002; 2002US-0411023P.		
PR	17-SEP-2002; 2002US-0411024P.		
PR	17-SEP-2002; 2002US-0411032P.		
PR	17-SEP-2002; 2002US-0411035P.		
PR	17-SEP-2002; 2002US-0411037P.		
PR	17-SEP-2002; 2002US-0411041P.		
PR	17-SEP-2002; 2002US-0411045P.		
PR	17-SEP-2002; 2002US-0411046P.		
PR	17-SEP-2002; 2002US-0411048P.		
PR	17-SEP-2002; 2002US-0411052P.		
PR	17-SEP-2002; 2002US-0411055P.		
PR	17-SEP-2002; 2002US-0411073P.		
PR	17-SEP-2002; 2002US-0411082P.		
PR	17-SEP-2002; 2002US-0411101P.		
PR	17-SEP-2002; 2002US-0411111P.		
PR	18-APR-2003; 2003US-0463700P.		
PR	18-APR-2003; 2003US-0463708P.		
PR	18-APR-2003; 2003US-0463716P.		
PR	18-APR-2003; 2003US-0463732P.		
PR	02-MAY-2003; 2003US-0467199P.		
PR	02-MAY-2003; 2003US-0467201P.		
PR	02-MAY-2003; 2003US-0467203P.		

```

PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472420P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Hatenbeck RF, Huang MM, Kothakota S, Haislan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2491; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytosstatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
XX Sequence 198 AA:
SQ
XX
XX Query Match 58.1%; Score 18; DB 8; Length 198;
XX Best Local Similarity 22.2%; Pred. No. 3.8e+02;
XX Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CXXXXXXXXX 12
Db 91 CATTTTTC 99
XX
XX RESULT 29
XX ADP30492
XX ID ADP30492 standard; protein; 198 AA.
XX
XX ADP30492;
XX
XX 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #1259.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM Cancer; Inflammatory; Immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX

```


PF 28-AUG-2003; 2003WO-US026780.
 XX
 PR 29-AUG-2002; 2002US-0406576P.
 PR 29-AUG-2002; 2002US-0406579P.
 PR 29-AUG-2002; 2002US-0406585P.
 PR 29-AUG-2002; 2002US-0406588P.
 PR 29-AUG-2002; 2002US-0406608P.
 PR 29-AUG-2002; 2002US-0406611P.
 PR 29-AUG-2002; 2002US-0406612P.
 PR 29-AUG-2002; 2002US-0406612P.
 PR 29-AUG-2002; 2002US-0406640P.
 PR 29-AUG-2002; 2002US-0406642P.
 PR 29-AUG-2002; 2002US-0406646P.
 PR 29-AUG-2002; 2002US-0406653P.
 PR 29-AUG-2002; 2002US-0406655P.
 PR 29-AUG-2002; 2002US-0406666P.
 PR 17-SEP-2002; 2002US-0410946P.
 PR 17-SEP-2002; 2002US-0410947P.
 PR 17-SEP-2002; 2002US-0410948P.
 PR 17-SEP-2002; 2002US-0410949P.
 PR 17-SEP-2002; 2002US-0410953P.
 PR 17-SEP-2002; 2002US-0410957P.
 PR 17-SEP-2002; 2002US-0410958P.
 PR 17-SEP-2002; 2002US-0410959P.
 PR 17-SEP-2002; 2002US-0410960P.
 PR 17-SEP-2002; 2002US-0410961P.
 PR 17-SEP-2002; 2002US-0410962P.
 PR 17-SEP-2002; 2002US-0411019P.
 PR 17-SEP-2002; 2002US-0411022P.
 PR 17-SEP-2002; 2002US-0411023P.
 PR 17-SEP-2002; 2002US-0411024P.
 PR 17-SEP-2002; 2002US-0411032P.
 PR 17-SEP-2002; 2002US-0411035P.
 PR 17-SEP-2002; 2002US-0411037P.
 PR 17-SEP-2002; 2002US-0411041P.
 PR 17-SEP-2002; 2002US-0411045P.
 PR 17-SEP-2002; 2002US-0411046P.
 PR 17-SEP-2002; 2002US-0411052P.
 PR 17-SEP-2002; 2002US-0411055P.
 PR 17-SEP-2002; 2002US-0411073P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411101P.
 PR 17-SEP-2002; 2002US-0411111P.
 PR 18-APR-2003; 2003US-0463700P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 02-MAY-2003; 2003US-0467230P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 19-MAY-2003; 2003US-0471336P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 22-MAY-2003; 2003US-0472430P.
 PR 09-JUN-2003; 2003US-0476609P.
 PR 09-JUN-2003; 2003US-0476641P.
 PR 08-JUL-2003; 2003US-0485218P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485224P.
 PR 08-JUL-2003; 2003US-0485325P.
 PR 14-JUL-2003; 2003US-0486446P.
 PR 14-JUL-2003; 2003US-0486480P.
 PR 15-JUL-2003; 2003US-0486891P.
 PR 15-JUL-2003; 2003US-0486906P.
 PR 08-AUG-2003; 2003US-0493341P.
 PR 08-AUG-2003; 2003US-0493370P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.
 XX
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
 XX

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
 PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linemann T;
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
 DR WPI; 2004-348438/32.
 XX
 PT New nucleic acid molecule for diagnosis, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 PT genetic, bacterial and viral diseases.
 XX
 PS Claim 1; SEQ ID NO 2490; 428bp; English.
 XX
 CC The present invention relates to an isolated nucleic acid molecule
 CC encoding a polypeptide which is believed to be cytostatic,
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is
 CC available on WIPOMB and is not in the specification.
 XX
 SQ Sequence 198 AA;
 Query Match 58.1%; Score 18; DB 8; Length 198;
 Best Local Similarity 22.2%; Pred. No. 3.8e+02;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 12
 Db 91 CATTYTTTC 99
 RESULT 30
 ADP0477
 ID ADP30477 standard; protein; 198 AA.
 AC XX
 ADP30477; XX
 DT 12-AUG-2004 (first entry)
 DE XX
 DE Human secreted protein SEQ ID #1244.
 XX
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
 KW cancer; inflammatory; immune; human secreted protein.
 OS Homo sapiens.
 XX
 PN WO2004035732-A2.
 PD XX
 PD 29-APR-2004.
 PF 28-AUG-2003; 2003WO-US026780.
 XX
 PR 29-AUG-2002; 2002US-0406576P.
 PR 29-AUG-2002; 2002US-0406579P.
 PR 29-AUG-2002; 2002US-0406585P.
 PR 29-AUG-2002; 2002US-0406588P.
 PR 29-AUG-2002; 2002US-0406608P.
 PR 29-AUG-2002; 2002US-0406611P.
 PR 29-AUG-2002; 2002US-0406612P.
 PR 29-AUG-2002; 2002US-0406612P.
 PR 29-AUG-2002; 2002US-0406640P.
 PR 29-AUG-2002; 2002US-0406642P.
 PR 29-AUG-2002; 2002US-0406646P.
 PR 29-AUG-2002; 2002US-0406653P.
 PR 29-AUG-2002; 2002US-0406655P.
 PR 29-AUG-2002; 2002US-0406666P.
 PR 17-SEP-2002; 2002US-0410946P.
 PR 17-SEP-2002; 2002US-0410947P.
 PR 17-SEP-2002; 2002US-0410948P.
 PR 17-SEP-2002; 2002US-0410949P.
 PR 17-SEP-2002; 2002US-0410953P.
 PR 17-SEP-2002; 2002US-0410957P.
 PR 17-SEP-2002; 2002US-0410958P.
 PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0410963P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486896P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halembeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX MPI; 2004-348436/32.
XX
XX
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX PT genetic, bacterial and viral diseases.
XX
XX
XX PS Claim 1; SEQ ID NO 2475; 429pp; English.
XX
XX
XX CC The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cytostatic,
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPWEB and is not in the specification.
XX
XX Sequence 198 AA;

Query Match 58.1%; Score 18; DB 8; Length 198;
Best Local Similarity 22.2%; Pred. No. 3.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 4 CXXXXXXC 12
Db 54 CATATTAC 62
RESULT 31
ADP30481
ID ADP30481 standard; protein; 198 AA.
XX
XX ADP30481;
XX
XX
XX 12-AUG-2004 (first entry)
XX
XX
XX DE Human secreted protein SEQ ID #1248.
XX
XX
XX KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX KW cancer; inflammatory; immune; human secreted protein.
XX
XX OS Homo sapiens.
XX
XX PN WO2004035732-A2.
XX
XX PD 29-APR-2004.
XX
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406665P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411111P.
XX 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485213P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486460P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.

PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

PS Claim 1; SEQ ID NO 2479; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and vincinical. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.

XX SQ Sequence 198 AA;

Query Match 58.1%; Score 18; DB 8; Length 198;
Best Local Similarity 22.2%; Pred. No. 3.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 54 CATTAATTAC 62

RESULT 32

ID ADP30491 standard; protein; 198 AA.

XX AC ADP30491;

DT 12-AUG-2004 (first entry)

XX Human secreted protein SEQ ID #1258.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.

XX OS Homo sapiens.
XX PN WO2004035732-A2.
XX PD 29-APR-2004.
XX PF 28-AUG-2003; 2003MO-US026780.

PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485213P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486460P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.

PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
DR
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2489; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
XX Sequence 198 AA;
SQ
Query Match 58.1%; Score 18; DB 8; Length 198;
Best Local Similarity 22.2%; Pred. No. 3.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 4 CXXXXXXXC 12
DB 91 CATTTC 99
RESULT 33
ADP31281
ID ADP31281 standard; protein; 228 AA.
XX
XX ADP31281;
AC
XX
XX 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #2048.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
OS
XX
XX WO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PF
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-048524P.
PR 08-JUL-2003; 2003US-048525P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
DR
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3279; 428bp; English.
PS
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.

SQ Sequence 228 AA;

Query Match 58.1%; Score 18; DB 8; Length 228;

Best Local Similarity 22.2%; Pred. No. 3.9e+02; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12

Db 183 CTTTAATC 191

RESULT 34

ABO80501

ID ABO80501 standard; protein; 229 AA.

XX ABO80501;

XX 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #12676.

KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

PN US651795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR MPI; 2003-615309/58.

DR N-PSDB; ABD14072.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 29247; 455bp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biotech technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

SQ Sequence 229 AA;

Query Match 58.1%; Score 18; DB 7; Length 229;

Best Local Similarity 22.2%; Pred. No. 3.9e+02; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12

Db 155 CATTATAC 163

RESULT 35

ABO77555

ID ABO77555 standard; protein; 233 AA.

XX ABO77555;

XX 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #9730.

KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

PN US651795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR MPI; 2003-615309/58.

DR N-PSDB; ABD11126.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 26301; 455bp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biotech technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

SQ Sequence 233 AA;

Query Match 58.1%; Score 18; DB 7; Length 233;

Best Local Similarity 22.2%; Pred. No. 3.9e+02; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12

Db 43 CSTASASC 51

RESULT 36

ADP31468
ID ADP31468 standard; protein; 234 AA.
XX
AC ADP31468;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2235.
XX
KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406589P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411010P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.

PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486466P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RP, Huang MM, Kothakota S, Hsieh L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
PS Claim 1, SEQ ID NO 3466; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic, the
CC antinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMER and is not in the specification.
XX
SQ Sequence 234 AA;
XX
Query Match 58.1%; Score 18; DB 8; Length 234;
Best Local Similarity 22.2%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CXXXXXXXXX 12
Db 213 CTTATTAC 221
RESULT 37
ADP31616
ID ADP31616 standard; protein; 242 AA.
XX
AC ADP31616;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2383.
XX
KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.

[illegible]

XX	18-FEB-1998;	98US-0074786P.
PR	27-JUL-1998;	98US-0094190P.
PA	(GENO-) GENOME THERAPEUTICS CORP.	
PI	Rubenfield MJ, Nolling J, Deloughery C, Bush D;	
DR	WPI; 2003-615309/58.	
N-PDB:	ABD06419.	
PT	Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,	
PP	pathological conditions resulting from bacterial infection.	
PS	Disclosure; SEQ ID NO 23594; 455pp; English.	
XX		
CC	The invention relates to Pseudomonas aeruginosa polypeptides and the	
CC	polynucleotides encoding them. The sequences are useful in diagnosis and	
CC	therapy of pathological conditions, as molecular targets for diagnostics,	
CC	prophylaxis and treatment of pathological conditions resulting from a	
CC	bacterial infection, for evaluating a compound, such as a polypeptide,	
CC	for the ability to bind a P. aeruginosa nucleic acid, as components of	
CC	effective antibacterial targets, as targets for antibacterial drugs,	
CC	including anti-P. aeruginosa drugs, as templates for recombinant	
CC	production of P. aeruginosa-derived peptides or polypeptides, as target	
CC	components for diagnosis and/or treatment of P. aeruginosa-caused	
CC	infection, and in detection of P. aeruginosa sequences or other sequences	
CC	of Pseudomonas species using biochip technology. Sequences AB067826-	
CC	AB084336 represent P. aeruginosa polypeptides of the invention. Note: The	
CC	sequence data for this patent did not form part of the printed	
CC	specification but was obtained in electronic format from USPTO at	
CC	seqdata.uspto.gov/sequence.html	
SQ	Sequence 246 AA;	
OY	4 CXXXXXXXC 12	
Db	202 CSTSSATC 210	
RESULT 40		
ID	ADP30754 standard; protein; 249 AA.	
XX	ADP30754;	
DT	12-AUG-2004 (first entry)	
XX	Human secreted protein SEQ ID #1521.	
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;	
OS	cancer; inflammatory; immune; human secreted protein.	
XX	Homo sapiens.	
PV	WO2004035732-A2.	
XX	29-APR-2004.	
XX	28-AUG-2003; 2003WO-US026780.	
XX	29-AUG-2002; 2002US-0406576P.	
XX	29-AUG-2002; 2002US-0406579P.	
XX	29-AUG-2002; 2002US-0406585P.	
XX	29-AUG-2002; 2002US-0406588P.	
XX	29-AUG-2002; 2002US-0406608P.	
XX	29-AUG-2002; 2002US-0406611P.	
XX	29-AUG-2002; 2002US-0406612P.	

PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411010P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halsebeck RF, Huang MM, Kochakota S, Haisnan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX MPI; 2004-348438/32.
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

XX
PS Claim 1; SEQ ID NO 2752; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 249 AA;
Query Match 58.1%; Score 18; DB 8; Length 249;
Best Local Similarity 22.2%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CXXXXXXXXC 12
DB 238 CTTAAAAAC 246
RESULT 41
ADP31396
ID ADP31396 standard; protein; 254 AA.
XX
AC ADP31396;
XX
DT 12-AUG-2004 (first entry)
XX
DB Human secreted protein SEQ ID #2163.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN MO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.

PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-047136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX PA
XX PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halembeck RP, Huang MM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX PT genetic, bacterial and viral diseases.
XX PS
XX PS Claim 1; SEQ ID NO 3394; 428bp; English.
XX
XX CC The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cytostatic,
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPOMB and is not in the specification.
XX
XX SQ Sequence 254 AA;

Query Match 58.1%; Score 18; DB 8; Length 254;
Best Local Similarity 22.2%; Pred. No. 4e+02; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 7;

QY 4 CXXXXXXXXC 12
| |
DB 77 CTATAAAC 85

RESULT 42
ADP30846
ID ADP30846 standard; protein; 279 AA.
XX
XX ADP30846;
XX
XX 12-AUG-2004 (first entry)
XX
XX DE Human secreted protein SEQ ID #1613.
XX
XX DE Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX KW cancer; inflammatory; immune; human secreted protein.
XX
XX OS Homo sapiens.
XX
XX PN MO2004035732-A2.
XX
XX PD 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406665P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411101P.
XX 17-SEP-2002; 2002US-0411111P.
XX 18-APR-2003; 2003US-0463700P.
XX 18-APR-2003; 2003US-0463708P.
XX 18-APR-2003; 2003US-0463716P.
XX 18-APR-2003; 2003US-0463732P.
XX 02-MAY-2003; 2003US-0467199P.
XX 02-MAY-2003; 2003US-0467201P.
XX 02-MAY-2003; 2003US-0467203P.
XX 02-MAY-2003; 2003US-0467230P.
XX 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486466P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisnan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
PS Claim 1; SEQ ID NO 2844; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 279 AA;
Query Match 58.1%; Score 18; DB 8; Length 279;
Best Local Similarity 22.2%; Pred. NO. 4.1e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 4 CXXXXXXXXX 12
DB 247 CTATTTAAC 255
RESULT 43
ADP31523
ID ADP31523 standard; protein; 279 AA.
XX
XX ADP31523;
XX
XX 12-AUG-2004 (first entry)
DE Human secreted protein SEQ ID #3290.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; Inflammatory; Immune; human secreted protein.
OS Homo sapiens.
XX
XX MO2004035732-A2.
PN
XX
PD 29-APR-2004.
XX
XX 28-AUG-2003; 2003MO-US026780.
XX

PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486466P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisnan L, Linnemann T;

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 3521; 428bp; English.
XX
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 279 AA;

Query Match 58.1%; Score 18; DB 8; Length 279;
Best Local Similarity 22.2%; Pred. No. 4,1e+02; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 7;

QY 4 CXXXXXXC 12
Db 66 CATTAAATC 74

RESULT 44
ADP30844
ID ADP30844 standard; protein; 279 AA.
XX
XX
AC ADP30844;
XX
XX
DT 12-AUG-2004 (first entry)
XX
XX
DE Human secreted protein SEQ ID #1611.
XX
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
XX
PN WO2004035732-A2.
XX
XX
PD 29-APR-2004.
XX
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406647P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0410969P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411025P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411047P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486896P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493373P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Haisan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX
XX WPI; 2004-348438/32.
XX
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 2842; 428bp; English.
XX
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 279 AA;

Query Match 58.1%; Score 18; DB 8; Length 279;

Best Local Similarity 22.2%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 247 CTATTTAAC 255

RESULT 45

AB48050
ID AAB48050 standard; protein; 281 AA.

AC AAB48050;

DT 19-MAR-2001 (first entry)

DE Signal transduction protein.

KM Zea mays; maize; signal transduction protein; phytohormone; ethylene;
auxin; cytokinin; gibberellin; immunogen.

OS Zea mays.

PN WO200070059-A2.

PD 23-NOV-2000.

PF 28-APR-2000; 2000MO-US011687.

PR 14-MAY-1999; 99US-0134292P.

PR 08-JUL-1999; 99US-0142996P.

PA (PION-) PIONEER HI-BRED INT INC.

PI Helentjaris TG;

DR WPI; 2001-031929/04.

DR N-PSDB; AAC84271.

PT New signal transduction nucleic acids and encoded proteins useful for
regulating phytohormone expression, including ethylene, auxin,
cytokinins and gibberellin, to provide control of plant response to
environmental stresses.

PS Claim 13; Page 102; 126pp; English.

CC The invention provides Zea mays signal transduction proteins and encoding
nucleotide sequences. The nucleic acids are useful for regulating
expression of phytohormones, including ethylene, auxins, cytokinins, and
gibberellin, to effect developmental changes in plants and provide
control of plant response to environmental stresses. They may also be
used as probes or amplification primers in the detection, quantitation or
isolation of gene transcripts, for detecting mutations in the gene, for
monitoring upregulation of expression or changes in enzyme activity in
screening assays of compounds, for detection of any number of allelic
variants, or for site-directed mutagenesis in eukaryotic cells. They may
further be used for recombinant expression of their encoded polypeptides,
as immunogens in the preparation or screening of antibodies, and in sense
or antisense suppression of genes in a host cell, tissue or plant. The
proteins may be used in assays for enzyme agonists or antagonists, as
immunogens or antigens to obtain antibodies specifically immunoreactive
with the proteins. The present sequence represents a signal transduction
protein of the invention

SQ Sequence 281 AA;

Query Match 58.1%; Score 18; DB 4; Length 281;

Best Local Similarity 22.2%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 11 CSATTSSC 19

RESULT 46...
ADM43729
ID ADM43729 standard; protein; 281 AA.

AC ADM43729;

DT 24-MAR-2005 (first entry)

DE Corn DREB-type transcription factor, seq id 2.

KM Drought resistance; crop improvement; bold tolerance;
transcription factor; DREB.

OS Zea mays.

PN CN1472222-A.

PD 04-FEB-2004.

PF 29-JUL-2002; 2002CN-00125372.

PR 29-JUL-2002; 2002CN-00125372.

PA (UYQI) UNIV QINGHUA.

PI Liu Q, Qin F, Zhao J;

DR WPI; 2004-317379/30.

DR N-PSDB; ADM43728.

PT DREB transcription factor of corn and its encoding genes and use.

PS Disclosure; SEQ ID NO 2; 18pp; Chinese.

CC The present invention discloses 2 corn DREB-type transcription factors,
CC MADREB1 and MADREB2, which are important in conditions of cold and
CC drought. Also disclosed are their coding genes and their application in
CC culture the cold and drought resistant plants. The current sequence
CC represents a DREB-type transcription factor.

SQ Sequence 281 AA;

Query Match 58.1%; Score 18; DB 8; Length 281;

Best Local Similarity 22.2%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 11 CSATTSSC 19

RESULT 47

AB079000
ID AB079000 standard; protein; 286 AA.

AC AB079000;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #1175.

KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

PN US6551795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

XX

PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
DR N-PSDB; ABD12571.
XX
PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 27746; 455bp; English.
XX
CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-*P. aeruginosa* drugs, as templates for recombinant
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology. Sequences AB067826-
CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 286 AA;

Query Match 58.1%; Score 18; DB 7; Length 286;
Best Local Similarity 22.2%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXC 12
DB 131 CSSAATAC 139

RESULT 48
ADP31192
ID ADP31192 standard; protein; 297 AA.
XX
AC ADP31192;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1959.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
XX

PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406656P.
PR 17-SEP-2002; 2002US-0410945P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493373P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir KJ, Beaurang PA, Behrens D;
PI Hahlenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX

PS Claim 1; SEQ ID NO 3190; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMER and is not in the specification.
SQ Sequence 297 AA;
XX
QY Query Match 58.1%; Score 18; DB 8; Length 297;
Best Local Similarity 22.2%; Pred. No. 4.2e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Db 39 CAAATTAC 47
4 CXXXXXXC 12
39 CAAATTAC 47
RESULT 49
ADP31454
ID ADP31454 standard; protein; 312 AA.
XX
AC ADP31454;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2221.
XX
KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN MO2004035732-AA2.
PD
XX 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.

PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-041101P.
PR 17-SEP-2002; 2002US-041111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467209P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-048524P.
PR 08-JUL-2003; 2003US-048525P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3452; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMER and is not in the specification.
SQ Sequence 312 AA;
XX
QY Query Match 58.1%; Score 18; DB 8; Length 312;
Best Local Similarity 22.2%; Pred. No. 4.2e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Db 7 CAAATTATTC 15
4 CXXXXXXC 12
7 CAAATTATTC 15

```
RESULT 50
ADP31615
ID ADP31615 standard; protein; 317 AA.
XX
AC ADP31615;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2382.
XX
KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX
PR 29-AUG-2002; 2002US-0406579P.
XX
PR 29-AUG-2002; 2002US-0406585P.
XX
PR 29-AUG-2002; 2002US-0406588P.
XX
PR 29-AUG-2002; 2002US-0406589P.
XX
PR 29-AUG-2002; 2002US-0406608P.
XX
PR 29-AUG-2002; 2002US-0406611P.
XX
PR 29-AUG-2002; 2002US-0406612P.
XX
PR 29-AUG-2002; 2002US-0406616P.
XX
PR 29-AUG-2002; 2002US-0406640P.
XX
PR 29-AUG-2002; 2002US-0406642P.
XX
PR 29-AUG-2002; 2002US-0406646P.
XX
PR 29-AUG-2002; 2002US-0406653P.
XX
PR 29-AUG-2002; 2002US-0406655P.
XX
PR 29-AUG-2002; 2002US-0406666P.
XX
PR 17-SEP-2002; 2002US-0410946P.
XX
PR 17-SEP-2002; 2002US-0410947P.
XX
PR 17-SEP-2002; 2002US-0410948P.
XX
PR 17-SEP-2002; 2002US-0410949P.
XX
PR 17-SEP-2002; 2002US-0410953P.
XX
PR 17-SEP-2002; 2002US-0410957P.
XX
PR 17-SEP-2002; 2002US-0410958P.
XX
PR 17-SEP-2002; 2002US-0410959P.
XX
PR 17-SEP-2002; 2002US-0410960P.
XX
PR 17-SEP-2002; 2002US-0410961P.
XX
PR 17-SEP-2002; 2002US-0410962P.
XX
PR 17-SEP-2002; 2002US-0410966P.
XX
PR 17-SEP-2002; 2002US-0411022P.
XX
PR 17-SEP-2002; 2002US-0411023P.
XX
PR 17-SEP-2002; 2002US-0411024P.
XX
PR 17-SEP-2002; 2002US-0411024P.
XX
PR 17-SEP-2002; 2002US-0411032P.
XX
PR 17-SEP-2002; 2002US-0411035P.
XX
PR 17-SEP-2002; 2002US-0411037P.
XX
PR 17-SEP-2002; 2002US-0411041P.
XX
PR 17-SEP-2002; 2002US-0411045P.
XX
PR 17-SEP-2002; 2002US-0411046P.
XX
PR 17-SEP-2002; 2002US-0411048P.
XX
PR 17-SEP-2002; 2002US-0411052P.
XX
PR 17-SEP-2002; 2002US-0411055P.
XX
PR 17-SEP-2002; 2002US-0411073P.
XX
PR 17-SEP-2002; 2002US-0411082P.
XX
PR 17-SEP-2002; 2002US-0411082P.
XX
PR 17-SEP-2002; 2002US-0411101P.
XX
PR 17-SEP-2002; 2002US-0411111P.
XX
PR 18-APR-2003; 2003US-0463700P.
XX
PR 18-APR-2003; 2003US-0463708P.
XX
PR 18-APR-2003; 2003US-0463716P.
XX
PR 18-APR-2003; 2003US-0463732P.
XX
PR 02-MAY-2003; 2003US-0467199P.
XX
PR 02-MAY-2003; 2003US-0467201P.
XX
PR 02-MAY-2003; 2003US-0467203P.
XX
PR 19-MAY-2003; 2003US-0471306P.
XX
PR 19-MAY-2003; 2003US-0471336P.
```

```
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR MPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3613; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 317 AA;
XX
QY Query Match 58.1%; Score 18; DB 8; Length 317;
Best Local Similarity 22.2%; Pred. No. 4.2e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CXXXXXXC 12
Db 145 CATTATATC 153
```

Search completed: January 4, 2006, 15:55:50
Job time : 115.652 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 ; Search time 15.5217 Seconds
(without alignments)
92.983 Million cell updates/sec

Title: US-09-932-322-3

Perfect score: 31
Sequence: 1 XXXXXXXXXXXXXXXX 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 1000 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	18	58.1	155	2 T28945	hypothetical prote
2	18	58.1	470	2 T45962	hypothetical prote
3	18	58.1	1237	2 A34598	ecdysone-induced p
4	18	58.1	1394	2 B34598	ecdysone-induced p
5	18	58.1	1443	2 S05979	steroid hormone re
6	17	54.8	199	2 T33355	hypothetical prote
7	17	54.8	212	2 S63342	hypothetical prote
8	17	54.8	214	2 C70812	probable lipoq prot
9	17	54.8	244	2 H72532	hypothetical prote
10	17	54.8	244	2 T40415	hypothetical prote
11	17	54.8	259	2 H71353	hypothetical prote
12	17	54.8	281	2 AC0661	conserved hypothet
13	17	54.8	293	2 T48975	probable lipoprote
14	17	54.8	298	2 T01948	xyloglucan endo-tr
15	17	54.8	315	2 A85019	hypothetical prote
16	17	54.8	373	2 AB1224	probable bHLH DNA-
17	17	54.8	373	2 AD1577	cobalamin biosynth
18	17	54.8	415	2 T43352	nuclear receptor N
19	17	54.8	507	2 T33024	hypothetical prote
20	17	54.8	530	2 T48004	multifunctional am
21	17	54.8	533	2 A42143	skn-1 - Caenorhabd
22	17	54.8	575	2 T07739	probable ferriochel
23	17	54.8	582	2 T07953	chromomodulin pre
24	17	54.8	582	2 T07953	lectin-like protei
25	17	54.8	582	2 T07953	lectin-like protei
26	17	54.8	584	2 B25682	homeotic protein B
27	17	54.8	589	2 S41011	hypothetical prote
28	17	54.8	595	2 A42086	CD30 antigen precu
29	17	54.8	600	2 D83286	hypothetical prote

30	17	54.8	617	2 T15408	hypothetical prote
31	17	54.8	624	2 A36962	laccase (EC 1.10.3
32	17	54.8	729	2 A41029	integrin beta-8 ch
33	17	54.8	791	2 A46140	diacylglycerol kin
34	17	54.8	796	2 B46140	diacylglycerol kin
35	17	54.8	799	2 A38308	integrin beta-5 ch
36	17	54.8	934	2 T05201	hypothetical prote
37	17	54.8	938	2 T41932	hypothetical prote
38	17	54.8	1995	2 T15881	hypothetical prote
39	17	54.8	2395	2 S50820	surface protein ty
40	17	54.8	2704	2 S09118	G surface protein
41	17	54.8	2718	2 A23475	G surface protein
42	17	54.8	8243	2 T31307	type I fatty acid
43	16	51.6	52	2 G82010	probable lipoprote
44	16	51.6	87	2 H86753	prophage pi2 prote
45	16	51.6	91	2 US0036	Clara cell 10K pro
46	16	51.6	96	1 UGMS	uteroglobin precu
47	16	51.6	96	2 A36581	polychlorinated bi
48	16	51.6	98	2 A10334	conserved hypothet
49	16	51.6	103	2 S76421	ferredoxin (Fp-25
50	16	51.6	104	2 J00863	hypothetical 11.6K
51	16	51.6	116	2 S09845	hypothetical prote
52	16	51.6	121	2 A44544	amine dehydrogenas
53	16	51.6	128	2 S74085	lutropin beta chal
54	16	51.6	138	2 T21792	hypothetical prote
55	16	51.6	148	2 B72513	hypothetical prote
56	16	51.6	152	2 H97405	hypothetical prote
57	16	51.6	153	2 P75478	conserved hypothet
58	16	51.6	156	2 T31839	hypothetical prote
59	16	51.6	158	2 A61091	lutropin beta chal
60	16	51.6	159	2 I51373	lutelinizing hormon
61	16	51.6	160	1 S17194	lutelinizing hormon
62	16	51.6	166	2 I51242	hypothetical prote
63	16	51.6	166	2 S6590	transcripton init
64	16	51.6	182	2 T01909	hypothetical prote
65	16	51.6	190	2 S48101	xyloglucan endo-1,
66	16	51.6	195	2 T19151	hypothetical prote
67	16	51.6	207	2 A69446	conserved hypothet
68	16	51.6	209	2 S51480	drought-induced pr
69	16	51.6	238	2 T04166	thamatin-like pro
70	16	51.6	246	2 AG3519	enterobactin synth
71	16	51.6	246	2 T20341	hypothetical prote
72	16	51.6	248	2 T21786	hypothetical prote
73	16	51.6	251	2 A55523	hypothetical prote
74	16	51.6	252	2 T04739	hypothetical prote
75	16	51.6	255	2 T38426	lymphocyte activat
76	16	51.6	260	2 T23033	hypothetical prote
77	16	51.6	261	2 C64049	hypothetical prote
78	16	51.6	261	2 S63604	homobox protein G
79	16	51.6	262	2 S01914	B1 membrane glycop
80	16	51.6	265	2 H97280	hydroxyethylthiaz
81	16	51.6	266	2 T02807	arsenate reductase
82	16	51.6	266	2 H84409	thiamin biosynthes
83	16	51.6	268	2 S19184	hypothetical prote
84	16	51.6	268	2 H81806	histidine-binding
85	16	51.6	268	2 B81063	amino acid ABC tra
86	16	51.6	269	2 T35647	phosphomethylpyrim
87	16	51.6	275	2 T32813	hypothetical prote
88	16	51.6	276	2 T01177	hypothetical prote
89	16	51.6	276	2 T52349	F-box protein FBX2
90	16	51.6	277	2 A46241	interferon respons
91	16	51.6	278	2 T319749	hypothetical prote
92	16	51.6	279	2 T51017	hypothetical prote
93	16	51.6	286	2 G97442	hypothetical prote
94	16	51.6	289	2 H75551	dihydropterolate sy
95	16	51.6	293	2 T31840	xyloglucan endo-1,
96	16	51.6	295	2 S48102	probable transcrip
97	16	51.6	299	2 F84785	hypothetical prote
98	16	51.6	303	2 T29321	hypothetical prote
99	16	51.6	303	2 T29321	hypothetical prote
100	16	51.6	327	2 A46484	apoptosis-mediatin
101	16	51.6	331	2 A71338	conserved hypothet
102	16	51.6	332	2 T26339	hypothetical prote

hypothetical prote
laccase (EC 1.10.3
integrin beta-8 ch
diacylglycerol kin
diacylglycerol kin
integrin beta-5 ch
hypothetical prote
hypothetical prote
hypothetical prote
surface protein ty
G surface protein
G surface protein
type I fatty acid
probable lipoprote
prophage pi2 prote
Clara cell 10K pro
uteroglobin precu
polychlorinated bi
conserved hypothet
ferredoxin (Fp-25
hypothetical 11.6K
hypothetical prote
amine dehydrogenas
lutropin beta chal
hypothetical prote
hypothetical prote
hypothetical prote
conserved hypothet
hypothetical prote
lutropin beta chal
lutelinizing hormon
lutelinizing hormon
hypothetical prote
transcripton init
hypothetical prote
xyloglucan endo-1,
hypothetical prote
conserved hypothet
drought-induced pr
thamatin-like pro
enterobactin synth
hypothetical prote
hypothetical prote
hypothetical prote
lymphocyte activat
hypothetical prote
hypothetical prote
homobox protein G
B1 membrane glycop
hydroxyethylthiaz
arsenate reductase
thiamin biosynthes
hypothetical prote
histidine-binding
amino acid ABC tra
phosphomethylpyrim
hypothetical prote
hypothetical prote
F-box protein FBX2
interferon respons
hypothetical prote
hypothetical prote
hypothetical prote
dihydropterolate sy
xyloglucan endo-1,
probable transcrip
hypothetical prote
apoptosis-mediatin
conserved hypothet
hypothetical prote

103	16	51.6	333	2	A96829	probable RING fing
104	16	51.6	336	2	S76979	hypothetical prote
105	16	51.6	341	2	S63666	platelet activatin
106	16	51.6	341	2	S43352	platelet-activatin
107	16	51.6	348	2	B75445	conserved hypotet
108	16	51.6	355	2	A34597	rhomboid protei
109	16	51.6	362	2	T36079	hypothetical prote
110	16	51.6	377	2	T05354	hypothetical prote
111	16	51.6	379	2	S14885	hypothetical prote
112	16	51.6	389	2	T29488	hypothetical prote
113	16	51.6	392	2	I54491	cell surface antig
114	16	51.6	416	1	JN0006	neve growth facto
115	16	51.6	418	2	T16713	hypothetical prote
116	16	51.6	429	2	T14861	gene G8G-154 prote
117	16	51.6	433	2	F86163	hypothetical prote
118	16	51.6	439	2	G97329	anaerobic dicarbox
119	16	51.6	468	2	T22397	hypothetical prote
120	16	51.6	480	2	T23608	hypothetical prote
121	16	51.6	483	2	T17346	hypothetical prote
122	16	51.6	491	2	T22844	hypothetical prote
123	16	51.6	492	2	S49147	BRF-2 protein - hu
124	16	51.6	496	2	S55665	conserved hypotet
125	16	51.6	521	2	F64522	hypothetical prote
126	16	51.6	523	2	B71985	hypothetical prote
127	16	51.6	525	2	G71317	hypothetical prote
128	16	51.6	528	2	D70968	hypothetical prote
129	16	51.6	543	2	F72071	conserved hypotet
130	16	51.6	543	2	C86551	CT384 hypothetical
131	16	51.6	545	2	H83342	periplasmic trehal
132	16	51.6	558	2	A36347	glypican 1 precurs
133	16	51.6	558	2	I56545	glypican precursor
134	16	51.6	561	2	I55235	rhoprty protein 2
135	16	51.6	579	2	T22703	hypothetical prote
136	16	51.6	583	2	JT0395	DNA-binding protei
137	16	51.6	596	2	S59944	hydroxymethylglut
138	16	51.6	604	2	T04120	hydroxymethylglut
139	16	51.6	618	2	S33044	hypothetical prote
140	16	51.6	625	2	S48941	regulatory protein
141	16	51.6	625	2	A48579	trophozoite surfac
142	16	51.6	667	2	T49226	hypothetical prote
143	16	51.6	687	2	C55926	DNA binding protei
144	16	51.6	692	2	A96502	hypothetical prote
145	16	51.6	713	2	A35502	major surface-labe
146	16	51.6	736	2	JC7518	subtilisin-like se
147	16	51.6	768	2	B41029	integrin beta-8 ch
148	16	51.6	773	1	A36932	iron(II) transport
149	16	51.6	773	2	C91160	ferrous iron trans
150	16	51.6	773	2	B86006	ferrous iron trans
151	16	51.6	797	2	A36811	hypothetical prote
152	16	51.6	812	1	PLMS	plasma(EC 3.4.21
153	16	51.6	829	2	T07406	probable protein k
154	16	51.6	864	2	S61148	SBE2 protein - yea
155	16	51.6	882	2	E96931	hypothetical prote
156	16	51.6	925	2	T02811	DNA excision/repai
157	16	51.6	939	2	C81730	metalloprotease,
158	16	51.6	942	2	B72015	metalloprotease,
159	16	51.6	942	2	C86610	insulinase family/
160	16	51.6	953	2	I48078	CHO1 antigen - Chi
161	16	51.6	962	2	JC5571	subtilisin-like pr
162	16	51.6	966	2	A33626	fibrinogen alpha c
163	16	51.6	969	1	A39490	subtilisin-like pr
164	16	51.6	975	2	JC5570	subtilisin-like pr
165	16	51.6	979	2	A89633	protein F13B6.6 [1
166	16	51.6	985	2	S67255	REVI protein - yea
167	16	51.6	997	2	A87320	TonB-dependent rec
168	16	51.6	1011	2	S65668	preproteins translo
169	16	51.6	1058	2	A53767	mucin MUC5B, trach
170	16	51.6	1058	2	T30178	mitotic checkpoint
171	16	51.6	1102	2	T31004	probable mitotic c
172	16	51.6	1111	2	T01239	hypothetical prote
173	16	51.6	1111	2	T26972	hypothetical prote
174	16	51.6	1151	2	H71347	hypothetical prote
175	16	51.6	1151	2	H71347	hypothetical prote
176	16	51.6	1176	2	T49482	hypothetical prote
177	16	51.6	1274	2	T42017	cysteine rich prot
178	16	51.6	1609	2	S25345	probable membrane
179	16	51.6	1729	2	A49282	fusion protein 1a/
180	16	51.6	1899	2	T49273	hypothetical prote
181	16	51.6	2105	2	T18968	probable serine-ty
182	16	51.6	2233	2	T28669	surface protei
183	15	48.4	25	1	SMR	metallothionein 51
184	15	48.4	25	1	S52359	hypothetical prote
185	15	48.4	43	2	D84641	hypothetical prote
186	15	48.4	44	2	T19161	hypothetical prote
187	15	48.4	46	1	VTVA1P	viscotoxin 1-PS -
188	15	48.4	46	1	VTVA1P	viscotoxin A2 - Bu
189	15	48.4	46	1	VTVA1P	viscotoxin B - Bu
190	15	48.4	52	2	JC1197	metallothionein II
191	15	48.4	59	1	B44203	ferredoxin [3Fe-4S
192	15	48.4	60	2	S44350	ferredoxin [4Fe-4S
193	15	48.4	61	2	SMHY1C	metallothionein I
194	15	48.4	61	2	I65234	testicular luteal
195	15	48.4	61	2	S00810	metallothionein IC
196	15	48.4	62	2	S15676	chorionic gonadotr
197	15	48.4	62	2	B55995	prostaglandin E2 r
198	15	48.4	67	2	T01958	hypothetical prote
199	15	48.4	77	2	I48725	Q300 protein - mou
200	15	48.4	78	2	S25619	hypothetical prote
201	15	48.4	78	2	D72764	probable ferredoxi
202	15	48.4	87	2	JT0623	Killer toxin HYI
203	15	48.4	88	2	JQ0514	cysteine-rich 9K p
204	15	48.4	88	2	S12126	cysteine-rich oute
205	15	48.4	88	2	S12125	cysteine-rich oute
206	15	48.4	88	2	E71513	probable 9kDa-cy
207	15	48.4	88	2	D81671	cysteine rich oute
208	15	48.4	92	1	KVRB38	Ig kappa chain V r
209	15	48.4	93	2	T10379	hypothetical prote
210	15	48.4	95	1	FEDH1	ferredoxin [2Fe-2S
211	15	48.4	95	1	FEDH2	ferredoxin [2Fe-2S
212	15	48.4	97	2	T03100	hypothetical prote
213	15	48.4	101	2	T48791	hypothetical prote
214	15	48.4	101	2	T51856	RIMH-H2 finger pro
215	15	48.4	103	2	H84599	hypothetical prote
216	15	48.4	104	1	KVRBXP	Ig kappa chain V r
217	15	48.4	104	2	S60405	hypothetical prote
218	15	48.4	107	1	KARB41	Ig kappa chain V r
219	15	48.4	107	2	A72782	hypothetical prote
220	15	48.4	108	1	KVRB12	Ig kappa chain V r
221	15	48.4	108	1	KVRB2K	Ig kappa chain V r
222	15	48.4	108	1	KVRB2K	Ig kappa chain V r
223	15	48.4	108	1	KVRB2K	Ig kappa chain V r
224	15	48.4	108	1	KVRB2K	Ig kappa chain V r
225	15	48.4	109	1	KVRB37	Ig kappa chain V r
226	15	48.4	109	1	KVRB37	Ig kappa chain V r
227	15	48.4	109	1	KVRB37	Ig kappa chain V r
228	15	48.4	110	1	KVRB29	Ig kappa chain V r
229	15	48.4	110	1	KVRB29	Ig kappa chain V r
230	15	48.4	110	1	KVRB29	Ig kappa chain V r
231	15	48.4	112	2	S21196	Ig kappa chain V r
232	15	48.4	112	2	S21196	Ig kappa chain V r
233	15	48.4	113	1	WNVNEM	hypothetical prote
234	15	48.4	113	2	PC1274	NS5 protein - hepa
235	15	48.4	113	2	PC1277	NS5 protein - hepa
236	15	48.4	113	2	PC1278	NS5 protein - hepa
237	15	48.4	113	2	PC1276	NS5 protein - hepa
238	15	48.4	113	2	PC1275	NS5 protein - hepa
239	15	48.4	114	2	S57692	probable membrane
240	15	48.4	117	1	K4RBF2	Ig kappa chain pre
241	15	48.4	117	1	BVECAR	arsenical resistan
242	15	48.4	117	2	G91131	probable truncated
243	15	48.4	118	2	PN0141	Ig kappa chain V r
244	15	48.4	118	2	PN0139	Ig kappa chain V r
245	15	48.4	118	2	B32536	Ig kappa chain V r
246	15	48.4	119	2	A61465	Ig kappa chain V r
247	15	48.4	120	2	D82022	truncated pili
248	15	48.4	121	2	S58424	seminal plasma pro
249	15	48.4	121	2	D69047	ychn protein homol

249	15	48.4	122	2	A27635	Ig heavy chain pre	322	15	48.4	189	2	T19559	hypothetical prote
250	15	48.4	122	2	C64965	hypothetical prote	323	15	48.4	190	2	C69345	hypothetical prote
251	15	48.4	122	2	B85825	probable structura	324	15	48.4	191	2	AH3005	conserved hypothet
252	15	48.4	122	2	B90979	probable structura	325	15	48.4	191	2	S69735	hypothetical prote
253	15	48.4	125	2	S53562	probable membrane	326	15	48.4	193	2	B95402	hypothetical prote
254	15	48.4	125	2	B70981	hypothetical prote	327	15	48.4	194	2	U01560	hypothetical 20.6k
255	15	48.4	125	2	T01628	hypothetical prote	328	15	48.4	195	2	T41796	hypothetical 20.6k
256	15	48.4	127	2	T44119	hypothetical prote	329	15	48.4	195	2	T07735	nodulin-20a - soyb
257	15	48.4	127	2	B83104	hypothetical prote	330	15	48.4	196	2	PN0478	vancomycin resista
258	15	48.4	127	2	C84302	hypothetical prote	331	15	48.4	197	2	A81877	probable membrane
259	15	48.4	128	2	T18072	hypothetical prote	332	15	48.4	197	2	F72628	hypothetical prote
260	15	48.4	132	2	B24255	chorion class A pr	333	15	48.4	204	2	T28939	hypothetical prote
261	15	48.4	132	2	PQ0394	genome polypeptid	334	15	48.4	205	2	S37804	hypothetical prote
262	15	48.4	132	2	PQ0396	genome polypeptid	335	15	48.4	206	2	UC7311	20K protei
263	15	48.4	134	2	D90229	hypothetical prote	336	15	48.4	206	2	T25139	hypothetical prote
264	15	48.4	135	2	S36197	hypothetical prote	337	15	48.4	208	2	S49196	Kunitz trypsin inh
265	15	48.4	138	2	S00512	lutropin beta chai	338	15	48.4	209	2	A83849	alpha-ribazole-5'-
266	15	48.4	138	2	A82022	truncated pili	339	15	48.4	209	2	S27494	nod - Bra
267	15	48.4	139	2	I52320	testicular lutein	340	15	48.4	209	2	E96517	hypothetical prote
268	15	48.4	139	2	S72900	permease c216b -	341	15	48.4	209	2	T29735	hypothetical prote
269	15	48.4	140	2	H84502	lutropin beta chai	342	15	48.4	211	2	C88103	protein M10G11.15
270	15	48.4	141	1	UTR0B	lutropin beta chai	343	15	48.4	212	2	T29906	hypothetical prote
271	15	48.4	141	1	UTR0B	lutropin beta chai	344	15	48.4	213	2	B25750	nodulin-26b - soyb
272	15	48.4	141	1	UTR0B	lutropin beta chai	345	15	48.4	213	2	A25750	nodulin-27 - soyb
273	15	48.4	141	1	UTR0B	lutropin beta chai	346	15	48.4	213	2	A55583	pyrogutamy1-pepti
274	15	48.4	141	2	JC4527	lutropin beta chai	347	15	48.4	214	2	D71282	probable deda prot
275	15	48.4	141	2	F82021	luteinizing hormon	348	15	48.4	215	2	C98278	hypothetical prote
276	15	48.4	142	2	JC6305	truncated pili	349	15	48.4	215	2	B57055	melanocortin-4 rec
277	15	48.4	143	2	S52595	neutrin precursor	350	15	48.4	215	2	A26669	nodulin-22 precurs
278	15	48.4	144	2	T50622	probable membrane	351	15	48.4	215	2	T24710	hypothetical prote
279	15	48.4	145	2	T42280	hypothetical prote	352	15	48.4	216	1	F69686	orotate phosphorib
280	15	48.4	145	2	T33037	hypothetical prote	353	15	48.4	216	2	S54430	fimriaral protein h
281	15	48.4	146	2	T01683	immobilization ant	354	15	48.4	217	2	S01358	salivary glue prot
282	15	48.4	150	2	T49563	hypothetical prote	355	15	48.4	220	2	S52005	nodulin-30 (Mpv30)
283	15	48.4	150	2	D87652	hypothetical prote	356	15	48.4	220	2	S42879	nodulin-30 - kidne
284	15	48.4	152	2	A32669	vasotocin 1 / neur	357	15	48.4	223	2	A23036	nodulin-23 - soybe
285	15	48.4	152	2	AG0060	probable exported	358	15	48.4	224	2	S07315	nodulin - soybean
286	15	48.4	156	2	H83361	Arsc protein PA227	359	15	48.4	225	2	S56809	probable membrane
287	15	48.4	158	2	S35787	US9 protein - bovi	360	15	48.4	225	1	WZBBP1	gene 58 protein -
288	15	48.4	159	2	B29879	vasotocin / neurop	361	15	48.4	225	2	D88492	protein T07E3.6 [1
289	15	48.4	160	2	T05385	hypothetical prote	362	15	48.4	225	2	D64598	hypothetical prote
290	15	48.4	161	1	S43893	H+-exporting ATPas	363	15	48.4	226	2	H96547	probable BHLH tran
291	15	48.4	161	2	S14480	arginine-vasotocin	364	15	48.4	228	2	B72655	hypothetical prote
292	15	48.4	161	2	T13688	destablase 2 homo	365	15	48.4	228	2	H71350	hypothetical prote
293	15	48.4	162	2	P97139	hypothetical prote	366	15	48.4	228	2	S26995	cellulose 1,4-beta
294	15	48.4	164	2	A41712	H+-exporting ATPas	367	15	48.4	229	2	S60454	glucose starvation
295	15	48.4	165	2	AH0215	4-hydroxyphenylace	368	15	48.4	230	2	S18989	fama protein - Str
296	15	48.4	165	2	T48806	hypothetical prote	369	15	48.4	231	2	T24669	hypothetical prote
297	15	48.4	169	1	KTH0B	choriogonadotropin	370	15	48.4	232	2	S76342	hypothetical prote
298	15	48.4	170	2	AE0630	4-hydroxyphenylace	371	15	48.4	235	2	F82754	cellulose 1,4-beta
299	15	48.4	170	2	A83135	4-hydroxyphenylace	372	15	48.4	236	2	S26993	regulatory protein
300	15	48.4	170	2	T06056	hypothetical prote	373	15	48.4	237	1	B42371	riboflavin-binding
301	15	48.4	172	1	B69138	hypothetical prote	374	15	48.4	238	1	VOCH	riboflavin-binding
302	15	48.4	173	2	B39141	transcription repr	375	15	48.4	238	1	LNRTMA	mannose-binding 1e
303	15	48.4	174	2	C69404	hypothetical prote	376	15	48.4	238	2	A64156	Mg2+-transporting
304	15	48.4	174	2	T49180	hypothetical prote	377	15	48.4	238	2	C83194	conserved hypothet
305	15	48.4	174	2	B85977	hypothetical prote	378	15	48.4	238	2	T16640	hypothetical prote
306	15	48.4	176	2	E71371	hypothetical prote	379	15	48.4	239	1	JC1373	ribonuclease (EC 3
307	15	48.4	177	2	S25551	thiamatin-like pro	380	15	48.4	241	2	T30761	hypothetical prote
308	15	48.4	177	2	T07843	probable lysine-ke	381	15	48.4	242	2	UC4360	riboflavin-binding
309	15	48.4	179	2	T19557	hypothetical prote	382	15	48.4	242	2	T29699	hypothetical prote
310	15	48.4	179	2	F97324	uncharacterized co	383	15	48.4	243	2	H69525	hydroxymethylpyrim
311	15	48.4	180	2	S15327	fibinairal protein M	384	15	48.4	244	2	T31838	hypothetical prote
312	15	48.4	180	2	C23308	nifQ protein - Rho	385	15	48.4	245	2	T24565	proteolipid protei
313	15	48.4	180	2	F88028	protein F46F5.15 (386	15	48.4	246	2	I51325	mannose-binding 1e
314	15	48.4	181	2	T04165	pathogenesis-relat	387	15	48.4	248	1	LNH0MC	mannose-binding 1e
315	15	48.4	182	2	T26109	hypothetical prote	388	15	48.4	248	2	S49323	chymotrypsin (EC 3
316	15	48.4	184	2	C72740	hypothetical prote	389	15	48.4	248	2	B72045	uridylylate kinase C
317	15	48.4	185	2	B69028	hypothetical prote	390	15	48.4	248	2	G86577	UMP kinase [import
318	15	48.4	185	2	B26669	probable phosphono	391	15	48.4	248	2	D96535	RING-H2 finger pro
319	15	48.4	186	1	DEPSNL	nodulin-20 precurs	392	15	48.4	249	2	T31837	hypothetical prote
320	15	48.4	187	2	T32826	amine dehydrogenas	393	15	48.4	250	2	T21785	hypothetical prote
321	15	48.4	189	2	S60587	non-structural prote	394	15	48.4	250	1	A49053	CD27 antigen precu

395	15	48.4	250	2	T31836	hypothetical prote	468	15	48.4	328	2	G69152	protein C24B5.5 [i
396	15	48.4	250	2	T31835	hypothetical prote	469	15	48.4	330	2	T25169	hypothetical prote
397	15	48.4	253	2	T49483	hypothetical prote	470	15	48.4	331	2	T20955	hypothetical prote
398	15	48.4	255	2	F69692	amino acid ABC tra	471	15	48.4	331	2	A54295	interferon alpha/b
399	15	48.4	256	2	B32393	T-cell antigen 4-1	472	15	48.4	331	2	S59501	interferon recepto
400	15	48.4	257	2	T28946	hypothetical prote	473	15	48.4	331	2	G96785	protein P10A5.24 [
401	15	48.4	258	2	G96798	hypothetical prote	474	15	48.4	332	2	A57055	melanocortin recep
402	15	48.4	259	2	T52619	TIN2-like protein	475	15	48.4	333	2	A47536	gene WNT3 protein
403	15	48.4	260	1	A46517	CD27 antigen precu	476	15	48.4	335	2	S52890	hypothetical prote
404	15	48.4	262	1	VG1HPC	E1 membrane glycop	477	15	48.4	335	2	A40036	apoptosis-mediatin
405	15	48.4	262	1	S47427	E1 membrane glycop	478	15	48.4	335	2	JE0115	zinc-finger protei
406	15	48.4	263	2	T30083	hypothetical prote	479	15	48.4	336	2	A46478	major surface anti
407	15	48.4	266	2	PQ0393	genome polypeptid	480	15	48.4	336	2	A30527	interferon gamma r
408	15	48.4	266	2	T36857	B2/7R protein - var	481	15	48.4	337	2	T38500	hypothetical prote
409	15	48.4	268	2	S08229	chlorophyll a/b-bi	482	15	48.4	339	2	H71265	GTP-binding regula
410	15	48.4	270	2	I64041	fchd protein - Hae	483	15	48.4	340	1	RGHUB3	G-protein beta-sub
411	15	48.4	272	2	T20991	hypothetical prote	484	15	48.4	340	2	I53871	DNA-binding protei
412	15	48.4	273	2	E75476	hypothetical prote	485	15	48.4	341	2	S72445	vancomycin resist
413	15	48.4	275	2	F69229	hypothetical prote	486	15	48.4	343	2	JN0249	38.7K hypothetical
414	15	48.4	276	2	AE0860	hypothetical prote	487	15	48.4	343	2	H65385	hypothetical prote
415	15	48.4	277	2	S38503	beta-lactamase (EC	488	15	48.4	344	2	F70922	NADH2 dehydrogen
416	15	48.4	277	2	S38671	beta-lactamase (EC	489	15	48.4	345	1	QXXL2M	surface antigen 51
417	15	48.4	277	2	I37552	OX40 homolog - hum	490	15	48.4	345	2	A05279	hypothetical prote
418	15	48.4	279	2	B64917	membrane permease,	491	15	48.4	346	2	T27896	probable Fe(II) tr
419	15	48.4	280	2	T02090	xyloglucan endo-1,	492	15	48.4	348	2	T06385	iron(III) dictytra
420	15	48.4	281	2	B64216	hypothetical prote	493	15	48.4	349	2	S74439	interferon recepto
421	15	48.4	281	2	T03509	probable arsenate	494	15	48.4	349	2	UC6311	hypothetical prote
422	15	48.4	285	2	G69204	ribose-phosphate p	495	15	48.4	350	2	S21400	wingless homolog X
423	15	48.4	285	2	T21962	hypothetical prote	496	15	48.4	352	2	A48828	Wnt-3A protein - m
424	15	48.4	286	2	T02119	probable DnaJ prot	497	15	48.4	352	2	A39532	conserved hypothet
425	15	48.4	289	1	MF1HPC	E1 membrane glycop	498	15	48.4	352	2	E95381	hypothetical prote
426	15	48.4	291	2	T28846	hypothetical prote	499	15	48.4	352	2	G71914	P2405.9 [imported]
427	15	48.4	292	2	D69176	conserved hypothet	500	15	48.4	353	2	C96711	L1 protein - rhesu
428	15	48.4	294	1	KFMG3	cisue factor prec	501	15	48.4	354	1	PIWLRI	Wnt-3 protein - mo
429	15	48.4	294	2	B84640	CONSTANS-like B-bo	502	15	48.4	355	2	A35503	CONSTANS-like 1 -
430	15	48.4	294	2	T23682	hypothetical prote	503	15	48.4	355	2	T51414	hypothetical prote
431	15	48.4	300	2	T05729	hypothetical prote	504	15	48.4	355	2	T29932	hypothetical prote
432	15	48.4	303	2	T46715	hypothetical prote	505	15	48.4	356	2	T20656	hypothetical prote
433	15	48.4	303	2	T04541	hypothetical prote	506	15	48.4	360	2	S54173	inorganic diposph
434	15	48.4	305	2	F82765	conserved hypothet	507	15	48.4	360	2	S09552	nodulin - soybean
435	15	48.4	305	2	T25117	hypothetical prote	508	15	48.4	361	2	T34595	ABC-transporter AT
436	15	48.4	307	2	T08125	CONSTANS protein h	509	15	48.4	361	2	T37938	hypothetical prote
437	15	48.4	307	2	D70741	hypothetical prote	510	15	48.4	362	2	S61924	cyst wall protein
438	15	48.4	307	2	T18846	hypothetical prote	511	15	48.4	363	2	B65104	hypothetical 38.5
439	15	48.4	312	2	A85778	probable flavoprot	512	15	48.4	364	2	G82595	ABC transporter ph
440	15	48.4	312	2	E90929	probable flavoprot	513	15	48.4	364	2	T30723	hypothetical prote
441	15	48.4	312	2	B64928	electron transfer	514	15	48.4	364	2	B85430	NS5 protein - hepa
442	15	48.4	312	2	F69842	3-oxoacyl- acyl-ca	515	15	48.4	365	2	JQ0879	probable alcohol d
443	15	48.4	312	2	B82839	biotin synthetas p	516	15	48.4	366	1	A30196	DNA-directed DNA p
444	15	48.4	312	2	T31834	hypothetical prote	517	15	48.4	366	1	A64107	phosphate transpor
445	15	48.4	312	2	T32379	hypothetical prote	518	15	48.4	366	2	T01169	hypothetical prote
446	15	48.4	314	2	I37383	FAS soluble protei	519	15	48.4	366	2	F72062	CT449 hypothetical
447	15	48.4	314	2	S55179	mRNA splice defec	520	15	48.4	366	2	A86561	prehnate dehydtra
448	15	48.4	314	2	G84336	hypothetical prote	521	15	48.4	368	2	S59565	probable adenosylm
449	15	48.4	317	2	S29204	melanotropin recep	522	15	48.4	369	2	T10750	comg3 protein - My
450	15	48.4	317	2	A64343	hypothetical prote	523	15	48.4	369	2	S73716	protein phosphatas
451	15	48.4	318	2	T28941	hypothetical prote	524	15	48.4	370	2	S54297	E2 protein - human
452	15	48.4	319	1	B42971	alcohol dehydrogen	525	15	48.4	372	1	W2WL31	alcohol dehydrogen
453	15	48.4	320	2	S75911	hypothetical prote	526	15	48.4	372	2	A63184	hypothetical prote
454	15	48.4	320	2	A53119	cell adhesion glyco	527	15	48.4	372	2	T23680	hypothetical prote
455	15	48.4	321	2	F84597	probable prolinae-r	528	15	48.4	374	2	T22149	hypothetical trans
456	15	48.4	321	2	D81018	iron(III) ABC tran	529	15	48.4	376	2	S70841	beta-lactamase (EC
457	15	48.4	321	2	B81962	probable membrane	530	15	48.4	377	1	QKXC	cysteine proteinas
458	15	48.4	321	2	T24773	hypothetical prote	531	15	48.4	377	1	T12042	beta-lactamase, pe
459	15	48.4	321	2	D96035	probable transposa	532	15	48.4	377	2	C86111	beta-lactamase [im
460	15	48.4	321	2	B96025	probable DNA-(apur	533	15	48.4	377	2	C91270	protein R09F10.7 [
461	15	48.4	323	2	T52563	probable RNA- (apur	534	15	48.4	380	2	B89588	hypothetical prote
462	15	48.4	324	2	UC2395	Fas antigen precur	535	15	48.4	380	2	H70590	hypothetical prote
463	15	48.4	327	2	D86474	probable RING zinc	536	15	48.4	381	2	T06696	lipase homolog T29
464	15	48.4	328	2	T19322	hypothetical prote	537	15	48.4	382	2	T52880	carbazole dioxygen
465	15	48.4	328	2	S01359	salivary glue prot	538	15	48.4	384	2	T46578	hypothetical prote
466	15	48.4	328	2	H64554	heat shock protein	539	15	48.4	385	2	T31493	hypothetical prote
467	15	48.4	328	2	H71954	probable lipid A b	540	15	48.4	389	2	T23167	hypothetical prote

541	15	48.4	390	2	D97305	NADH-dependent but	614	15	48.4	485	2	T13479	hypothetical prote
542	15	48.4	390	2	B47013	butanol dehydrogen	615	15	48.4	487	2	T19875	hypothetical prote
543	15	48.4	392	2	C82628	hypothetical prote	616	15	48.4	490	2	T32003	hypothetical prote
544	15	48.4	392	2	T16776	hypothetical prote	617	15	48.4	493	2	S36931	chitinase (EC 3.2.
545	15	48.4	393	2	T16046	hypothetical prote	618	15	48.4	498	2	T46203	interleukin-14 pre
546	15	48.4	393	2	S48288	probable phosphopr	619	15	48.4	500	2	A71306	hypothetical prote
547	15	48.4	394	2	H72505	hypothetical prote	620	15	48.4	500	2	C84771	probable glucosyl
548	15	48.4	394	2	F95973	probable sugar upt	621	15	48.4	501	2	S43536	T-cell surface gly
549	15	48.4	396	1	TRXXB4	alpha-lytic protei	622	15	48.4	501	2	JH0447	alpha-1A-adrenerg
550	15	48.4	396	1	DCBYDM	adenosylmethionine	623	15	48.4	501	2	E96758	hypothetical prote
551	15	48.4	396	1	E64987	bicyclomycin resis	624	15	48.4	504	1	B48939	cellulose 1,4-beta
552	15	48.4	396	2	B91013	bicyclomycin resis	625	15	48.4	507	2	S18450	variant surface gl
553	15	48.4	396	2	D85857	bicyclomycin resis	626	15	48.4	509	1	S14629	aldehyde dehydrog
554	15	48.4	398	2	F40511	hypothetical prote	627	15	48.4	510	2	S41943	cellulose 1,4-beta
555	15	48.4	398	2	S17428	interleukin-1 rece	628	15	48.4	510	2	A70580	cellulose 1,4-beta
556	15	48.4	400	2	JC4591	alpha-1,3 fucosyl	629	15	48.4	511	2	S44716	cellulose 1,4-beta
557	15	48.4	400	2	T32705	hypothetical prote	630	15	48.4	512	2	A87270	hypothetical prote
558	15	48.4	405	2	B36340	alpha(1,3)-fucosyl	631	15	48.4	513	2	T37612	cytosol aminopept
559	15	48.4	406	2	T01931	adenyl cyclase -	632	15	48.4	515	2	A40491	alpha-1-adrenerg
560	15	48.4	407	2	JE0113	zinc-finger protei	633	15	48.4	515	2	JC1525	alpha-1B-adrenerg
561	15	48.4	410	2	JC7757	function-unknown p	634	15	48.4	515	2	I39073	interferon alpha-b
562	15	48.4	411	2	T29475	hypothetical prote	635	15	48.4	516	2	S33164	cellulose 1,4-beta
563	15	48.4	414	2	H84596	hypothetical prote	636	15	48.4	516	2	JS0083	cellulose 1,4-beta
564	15	48.4	416	2	I45890	ferrochelatase (EC	637	15	48.4	517	2	A45121	alpha-1B adrenerg
565	15	48.4	417	2	B86919	probable seeryl-tRN	638	15	48.4	520	1	B86299	hypothetical prote
566	15	48.4	417	2	E30341	alpha-1-adrenergic	639	15	48.4	520	1	RD1LNTS	dihydrofolate redu
567	15	48.4	417	2	T20199	hypothetical prote	640	15	48.4	520	1	RD1LNTZ	hypothetical prote
568	15	48.4	418	2	E71716	proline/betaine tr	641	15	48.4	520	2	T20007	hypothetical prote
569	15	48.4	423	2	A36794	hypothetical prote	642	15	48.4	520	2	S14600	E2 glycoprotein pr
570	15	48.4	423	2	T19581	hypothetical prote	643	15	48.4	521	2	E82508	methy1-accepting c
571	15	48.4	425	2	D88115	protein F53C3.11 f	644	15	48.4	526	2	B81679	conserved hypothet
572	15	48.4	425	2	C89753	protein F11C7.3 f1	645	15	48.4	529	2	S62194	hypothetical prote
573	15	48.4	427	2	G88492	protein T07E3.4 f1	646	15	48.4	533	2	A45392	RNA-directed RNA p
574	15	48.4	429	1	T03607	S-locus-specific g	647	15	48.4	538	2	B56110	tyrosine phosphopr
575	15	48.4	431	2	T14414	dicarboxylic acid	648	15	48.4	538	2	AB3632	D-ribulokinase (EC
576	15	48.4	432	2	A83616	hypothetical prote	649	15	48.4	539	2	D71522	hypothetical prote
577	15	48.4	432	2	T34154	dermal glycoprotei	650	15	48.4	540	2	S41942	cellulose 1,4-beta
578	15	48.4	433	2	T14329	alpha-1,3-fucosyl	651	15	48.4	540	2	B64829	membrane protein b
579	15	48.4	433	2	A57596	hypothetical prote	652	15	48.4	540	2	H90751	probable transport
580	15	48.4	434	2	G86163	hypothetical prote	653	15	48.4	540	2	F85615	probable transport
581	15	48.4	436	1	S49458	diphosphate-fructo	654	15	48.4	541	2	D84681	hypothetical prote
582	15	48.4	436	1	D69270	ferredoxin-nitrite	655	15	48.4	550	1	VG1HD6	E2 glycoprotein pr
583	15	48.4	438	1	ERADC2	early E1B 49K prot	656	15	48.4	552	2	S47034	cell fusion glycop
584	15	48.4	439	2	T15748	hypothetical prote	657	15	48.4	553	1	I46329	cell fusion glycop
585	15	48.4	445	2	A48881	serotonin receptor	658	15	48.4	553	1	VGXZGB	cell fusion glycop
586	15	48.4	445	2	T21744	hypothetical prote	659	15	48.4	553	1	A36830	cell fusion glycop
587	15	48.4	448	2	S36402	serotonin receptor	660	15	48.4	553	1	A46329	cell fusion glycop
588	15	48.4	448	2	A47519	serotonin receptor	661	15	48.4	553	1	B36830	cell fusion glycop
589	15	48.4	449	1	A42800	beta-alanine-pyruv	662	15	48.4	553	1	B46329	cell fusion glycop
590	15	48.4	450	2	E96695	hypothetical prote	663	15	48.4	553	1	D46329	cell fusion glycop
591	15	48.4	450	2	B36810	hypothetical prote	664	15	48.4	553	1	B46329	cell fusion glycop
592	15	48.4	451	2	T23731	hypothetical prote	665	15	48.4	553	1	G46329	cell fusion glycop
593	15	48.4	454	2	T26654	hypothetical prote	666	15	48.4	553	1	VGXZND	cell fusion glycop
594	15	48.4	456	2	T16796	hypothetical prote	667	15	48.4	553	1	VGXZNV	cell fusion glycop
595	15	48.4	458	2	B83341	L-serine dehydrata	668	15	48.4	553	1	VGXZTE	cell fusion glycop
596	15	48.4	458	2	T16123	hypothetical prote	669	15	48.4	553	1	VGXZU1	cell fusion glycop
597	15	48.4	461	2	C82399	conserved hypothet	670	15	48.4	553	2	S06345	cell fusion glycop
598	15	48.4	461	2	A35356	tumor necrosis fac	671	15	48.4	553	2	S38784	gene F protein - N
599	15	48.4	463	1	T43573	hypothetical prote	672	15	48.4	553	2	S40163	cell fusion protei
600	15	48.4	469	1	NM1V27	exo-alpha-sialidas	673	15	48.4	553	2	S38786	gene F protein - N
601	15	48.4	469	1	NM1V2	exo-alpha-sialidas	674	15	48.4	553	2	S23621	gene F protein - N
602	15	48.4	469	1	NM1V2	exo-alpha-sialidas	675	15	48.4	553	2	S23620	gene F protein - N
603	15	48.4	469	1	J01644	exo-alpha-sialidas	676	15	48.4	553	2	S38785	gene F protein - N
604	15	48.4	470	2	S78440	phosphoglucumutase	677	15	48.4	553	2	T15094	hypothetical prote
605	15	48.4	471	2	S36655	UDP rhamnose-antio	678	15	48.4	553	2	S23622	gene F protein - N
606	15	48.4	473	2	S60290	anthocyanin rhamo	679	15	48.4	554	2	B82934	hypothetical prote
607	15	48.4	473	2	AG0612	probable transport	680	15	48.4	554	2	A31300	perforin precursor
608	15	48.4	474	1	VYHUD	vitamin D-binding	681	15	48.4	554	2	A45818	cytochrome precurs
609	15	48.4	476	2	S09152	translation elonga	682	15	48.4	555	2	A37181	perforin 1 precurs
610	15	48.4	477	2	S71323	alpha-1A adrenerg	683	15	48.4	555	2	S63137	hypothetical prote
611	15	48.4	481	2	A89102	protein F25E5.4 f1	684	15	48.4	557	2	A48434	variant-specific g
612	15	48.4	483	2	B84454	hypothetical prote	685	15	48.4	560	2	A38731	alpha-1A adrenerg
613	15	48.4	483	2	T20895	hypothetical prote	686	15	48.4	563	2	C86397	protein TYN9.10 f1

687	15	48.4	567	2	D83400	conserved hypothet	760	15	48.4	737	2	AH1958	hypothetical prote
688	15	48.4	568	2	E96648	hypothetical prote	761	15	48.4	739	2	T12964	subtilisin homolog
689	15	48.4	571	2	S69210	protein kinase cak	762	15	48.4	740	2	T03847	fse-binding protei
690	15	48.4	572	2	I39369	alpha-1A-adrenergi	763	15	48.4	753	2	B36268	platelet glycoprot
691	15	48.4	573	2	T29880	hypothetical prote	764	15	48.4	753	2	T28787	hypothetical prote
692	15	48.4	572	2	T25397	hypothetical prote	765	15	48.4	761	2	JC5759	brain-specific ser
693	15	48.4	580	2	S72211	N-acetyl-beta-D-gl	766	15	48.4	766	2	S61424	inorganic diphosph
694	15	48.4	581	2	S17150	potassium channel	767	15	48.4	770	1	A38230	inorganic diphosph
695	15	48.4	590	2	D84523	probable auxin-reg	768	15	48.4	770	2	T00204	LDL receptor relat
696	15	48.4	590	2	A45283	interferon alpha/b	769	15	48.4	772	2	D96504	protein PgC16.25 f
697	15	48.4	590	2	A48461	ovarian abundant m	770	15	48.4	772	2	A60798	platelet glycoprot
698	15	48.4	597	2	J00107	hypothetical 66K p	771	15	48.4	778	2	T03156	ribonucleoside-dip
699	15	48.4	600	2	A47391	serum albumin prec	772	15	48.4	780	2	T21276	integrin beta3 - C
700	15	48.4	603	2	S28941	coagulation factor	773	15	48.4	781	2	S43534	hypothetical 82K p
701	15	48.4	605	1	ABPGS	serum albumin prec	774	15	48.4	784	2	U00317	platelet glycoprot
702	15	48.4	607	2	T34040	hypothetical prote	775	15	48.4	788	2	A26547	platelet glycoprot
703	15	48.4	608	1	ABONS1	serum albumin 1 pr	776	15	48.4	788	2	I17349	integrin beta-3 su
704	15	48.4	608	1	ABONS2	serum albumin 2 pr	777	15	48.4	788	2	I51530	hypothetical prote
705	15	48.4	608	1	ABRNS	serum albumin prec	778	15	48.4	792	2	T21276	hypothetical prote
706	15	48.4	608	1	S57632	serum albumin prec	779	15	48.4	796	2	H82406	probable lipase VC
707	15	48.4	608	2	T19935	hypothetical prote	780	15	48.4	796	2	UC7355	peroxisome prolif
708	15	48.4	609	1	ABHUS	serum albumin prec	781	15	48.4	796	2	T16424	F-spondin precurs
709	15	48.4	610	2	S35049	mucin JBR57 - huma	782	15	48.4	803	2	A47723	hypothetical prote
710	15	48.4	611	2	D96607	hypothetical prote	783	15	48.4	811	2	T25166	hypothetical prote
711	15	48.4	612	2	T05750	auxin-regulated pr	784	15	48.4	813	2	T21192	DNA polymerase III
712	15	48.4	612	2	S55084	probable membrane	785	15	48.4	815	2	T05555	hypothetical prote
713	15	48.4	613	2	A39402	potassium channel	786	15	48.4	818	2	T32154	hypothetical prote
714	15	48.4	613	2	T00853	hypothetical prote	787	15	48.4	819	2	C84615	hypothetical prote
715	15	48.4	614	2	B84006	formate dehydrogen	788	15	48.4	822	2	T01095	hypothetical prote
716	15	48.4	618	2	T00476	probable vacuolar	789	15	48.4	825	2	T29634	hypothetical prote
717	15	48.4	619	1	JH0776	hydrogenase (EC 1.	790	15	48.4	827	1	S28273	diacylglycerol kin
718	15	48.4	619	2	T08613	hypothetical prote	791	15	48.4	828	2	S52393	beta-galactosidase
719	15	48.4	620	2	T27008	hypothetical prote	792	15	48.4	833	2	T10695	transcription fact
720	15	48.4	621	2	S72493	laccase (EC 1.10.3	793	15	48.4	836	2	T28250	ORF MSV089 probab
721	15	48.4	621	2	JC7278	adaptor protein co	794	15	48.4	836	2	E96561	hypothetical prote
722	15	48.4	621	2	T22804	hydrogenase (EC 1.	795	15	48.4	840	2	H86429	hypothetical prote
723	15	48.4	624	2	S22703	hypothetical prote	796	15	48.4	842	2	T05400	hypothetical prote
724	15	48.4	624	2	G87619	voltage-gated pota	797	15	48.4	851	2	S12159	env protein - huma
725	15	48.4	626	2	T04895	sensor histidine k	798	15	48.4	852	1	GNNYHA	genome polypept
726	15	48.4	627	2	H82573	vacuolar sorting r	799	15	48.4	852	1	VCLJGG	env polypept
727	15	48.4	627	2	B48442	membrane transport	800	15	48.4	852	1	T01364	env polypept
728	15	48.4	627	4	A40201	arilfact-warning s	801	15	48.4	858	1	VCLJG2	env polypept
729	15	48.4	629	2	S60385	probable membrane	802	15	48.4	859	1	VCLJST	env polypept
730	15	48.4	630	2	T43460	hypothetical prote	803	15	48.4	859	1	VCLJCT	env polypept
731	15	48.4	631	2	A54659	DNA repair protein	804	15	48.4	859	1	VCLJ22	env polypept
732	15	48.4	631	2	T31782	hypothetical prote	805	15	48.4	859	1	VCLJBE	env polypept
733	15	48.4	631	2	T29926	hypothetical prote	806	15	48.4	859	1	VCLJBE	env polypept
734	15	48.4	632	2	S42731	collagen alpha 1 c	807	15	48.4	859	1	VCLJEB	env polypept
735	15	48.4	632	2	A71259	probable dicarboxy	808	15	48.4	859	1	VCLJEV	env polypept
736	15	48.4	633	2	A36353	DNA repair protein	809	15	48.4	859	1	VCLJEW	env polypept
737	15	48.4	633	2	I64143	hypothetical prote	810	15	48.4	859	1	VCLJMS	env polypept
738	15	48.4	640	2	S69546	phosphoenolpyruvat	811	15	48.4	859	2	S24571	env protein - huma
739	15	48.4	653	2	H86373	protein T23E23.16	812	15	48.4	860	2	C96504	protein PgC16.23 f
740	15	48.4	654	2	B75587	probable N-glycosyl	813	15	48.4	860	2	T37768	probable vacuolar
741	15	48.4	663	2	B65025	hypothetical prote	814	15	48.4	861	2	A48825	Notch homolog Motc
742	15	48.4	663	2	A91048	hypothetical prote	815	15	48.4	862	2	S64821	probable membrane
743	15	48.4	664	2	T06598	dnak-type molecula	816	15	48.4	863	2	D88465	protein B0244.7 f1
744	15	48.4	665	1	A42792	succinate dehydrog	817	15	48.4	864	2	B86449	hypothetical prote
745	15	48.4	670	2	T32221	hypothetical prote	818	15	48.4	869	2	S53098	envelope polypept
746	15	48.4	678	2	H96552	hypothetical prote	819	15	48.4	870	2	A96637	hypothetical prote
747	15	48.4	679	2	B85892	hypothetical prote	820	15	48.4	884	2	T18649	hypothetical prote
748	15	48.4	692	2	S37159	NADPH-ferrihemopro	821	15	48.4	890	2	F84548	hypothetical prote
749	15	48.4	695	2	S49163	transferrin precur	822	15	48.4	892	2	A41697	nitrate assimilat
750	15	48.4	699	2	A43734	probable protein k	823	15	48.4	909	2	H87729	protein Y23H5A.7 f
751	15	48.4	703	2	B82148	ATP-dependent heli	824	15	48.4	915	2	T03589	probable aspartate
752	15	48.4	704	2	B84685	hypothetical prote	825	15	48.4	915	2	T16623	hypothetical prote
753	15	48.4	704	2	T38117	probable protein k	826	15	48.4	923	2	E83574	hypothetical prote
754	15	48.4	708	2	T19474	hypothetical prote	827	15	48.4	925	2	T37475	lipoprotein recept
755	15	48.4	711	2	T30107	hypothetical prote	828	15	48.4	929	2	T32492	hypothetical prote
756	15	48.4	712	1	VCLJSD4	env polypept	829	15	48.4	931	2	T33744	hypothetical prote
757	15	48.4	720	2	A36526	cholesterol O-acetyltr	830	15	48.4	942	2	C96574	hypothetical prote
758	15	48.4	728	2	T20561	hypothetical prote	831	15	48.4	946	2	T00024	ent-kaurene synth
759	15	48.4	736	2	T12963	subtilisin homolog	832	15	48.4	959	2	E85276	hypothetical prote

833	15	48.4	960	2	T00808	hypothetical prote
834	15	48.4	964	1	VCLJC6	env polypepten pr
835	15	48.4	964	2	S13329	hypothetical prote
836	15	48.4	966	1	VCLJJC	env polypepten pr
837	15	48.4	979	1	JC2349	protein-tyrosine-p
838	15	48.4	979	2	D96574	hypothetical prote
839	15	48.4	982	1	VCLJVS	env polypepten pr
840	15	48.4	982	1	VCLJLK	env polypepten -
841	15	48.4	983	1	E45390	env polypepten pr
842	15	48.4	985	1	VCLJSP	insulin II gene en
843	15	48.4	989	2	T48845	env polypepten pr
844	15	48.4	990	1	G46535	env polypepten pr
845	15	48.4	993	2	A47500	Ig mu chain switch
846	15	48.4	993	2	S35633	DNA-binding protei
847	15	48.4	994	2	S18739	env protein - siml
848	15	48.4	996	2	T48721	prp 35 protein - m
849	15	48.4	1011	2	T17430	tol protein - Neur
850	15	48.4	1014	2	P96501	hypothetical prote
851	15	48.4	1016	2	T05066	hypothetical prote
852	15	48.4	1022	2	T48358	hypothetical prote
853	15	48.4	1025	2	T45811	Ca2+-transporting
854	15	48.4	1025	2	T14453	Ca2+-transporting
855	15	48.4	1030	2	T00812	Ca2+-transporting
856	15	48.4	1032	2	T43257	beta-1,3 exoglucan
857	15	48.4	1042	2	T16169	hypothetical prote
858	15	48.4	1042	2	T29307	hypothetical prote
859	15	48.4	1044	2	E86613	ribonucleoside red
860	15	48.4	1044	2	A72010	ribonucleoside-dip
861	15	48.4	1047	2	F81728	ribonucleoside-dip
862	15	48.4	1053	2	D71466	probable ribonucle
863	15	48.4	1054	2	T30177	cytoskeleton assem
864	15	48.4	1074	2	JC5928	semaphorin F precu
865	15	48.4	1076	2	B96682	protein FIR22.14 l
866	15	48.4	1082	2	T44177	hypothetical prote
867	15	48.4	1082	2	T43590	hypothetical prote
868	15	48.4	1086	2	T09325	probable capsid as
869	15	48.4	1086	2	T33893	hypothetical prote
870	15	48.4	1110	1	B42544	G2-G1 polypepten
871	15	48.4	1128	2	AD3008	peptide synthetase
872	15	48.4	1137	2	A98276	hypothetical prote
873	15	48.4	1137	2	S65440	nitric-oxide synth
874	15	48.4	1162	2	B97852	hypothetical prote
875	15	48.4	1168	2	I56985	kallinin B1 - mouse
876	15	48.4	1169	2	C71639	hypothetical prote
877	15	48.4	1170	2	A53612	laminin B1k chain
878	15	48.4	1172	1	TSHUP2	thrombospondin 2 p
879	15	48.4	1172	2	A42587	thrombospondin 2 p
880	15	48.4	1184	1	A39804	thrombospondin pre
881	15	48.4	1184	2	T46039	hypothetical prote
882	15	48.4	1192	2	S69000	laminin gamma 2 ch
883	15	48.4	1205	2	T41987	hypothetical prote
884	15	48.4	1217	2	T00270	hypothetical prote
885	15	48.4	1230	2	T42735	TBP-interacting pr
886	15	48.4	1232	2	T43027	neural cell adhesi
887	15	48.4	1260	2	T01334	hypothetical prote
888	15	48.4	1292	2	T09229	galactose binding
889	15	48.4	1299	2	T43251	furin (EC 3.4.21.7
890	15	48.4	1315	2	T50262	probable nucleopor
891	15	48.4	1319	2	A28513	glued protein - fr
892	15	48.4	1353	2	JC4279	adenylate cyclase
893	15	48.4	1358	2	A03905	genome polypepten
894	15	48.4	1386	2	T00257	hypothetical prote
895	15	48.4	1391	2	S50608	hypothetical prote
896	15	48.4	1391	2	T20406	hypothetical prote
897	15	48.4	1414	1	S23809	collagen alpha 2(I
898	15	48.4	1414	2	T33236	hypothetical prote
899	15	48.4	1416	2	E88550	protein ZC84.1 [lm
900	15	48.4	1437	2	S07430	M polypepten prec
901	15	48.4	1437	2	P96783	unknown protein F2
902	15	48.4	1453	2	A36861	orf 1b protein - L
903	15	48.4	1500	2	T03824	probable immediate
904	15	48.4	1502	2	S45429	probable membrane
905	15	48.4	1513	2	T23681	hypothetical prote
906	15	48.4	1587	2	G86467	hypothetical prote
907	15	48.4	1599	2	T16210	hypothetical prote
908	15	48.4	1732	2	G84664	hypothetical prote
909	15	48.4	1753	2	T00350	hypothetical prote
910	15	48.4	1755	2	T51532	hypothetical prote
911	15	48.4	1765	2	T42388	sodium channel alp
912	15	48.4	1773	2	T05128	hypothetical prote
913	15	48.4	1817	2	T34249	hypothetical prote
914	15	48.4	1820	2	A55494	latent transformin
915	15	48.4	1896	2	B72175	D1SR protein - var
916	15	48.4	1897	2	T28621	hypothetical prote
917	15	48.4	1965	2	T33216	hypothetical prote
918	15	48.4	2103	1	JQ1621	genome polypepten
919	15	48.4	2115	2	S38480	nonstructural prot
920	15	48.4	2120	2	T30243	alpha tectorin - c
921	15	48.4	2133	2	T30637	hypothetical prote
922	15	48.4	2149	2	C96695	ribulose biphosph
923	15	48.4	2155	2	T30197	alpha tectorin - m
924	15	48.4	2180	2	T29764	hypothetical prote
925	15	48.4	2205	1	MNWVRN	nonstructural poly
926	15	48.4	2214	1	OZBYU2	pyrimidine synthes
927	15	48.4	2222	1	A36028	DNA-directed DNA p
928	15	48.4	2227	1	GNNYHB	genome polypepten
929	15	48.4	2227	1	GNNYHM	genome polypepten
930	15	48.4	2227	1	GNNYHR	genome polypepten
931	15	48.4	2227	1	GNNYMK	genome polypepten
932	15	48.4	2230	1	GNNYSA	genome polypepten
933	15	48.4	2254	2	T09053	low voltage-activa
934	15	48.4	2276	2	T00076	hypothetical prote
935	15	48.4	2476	2	T34022	zonadhesin - pig
936	15	48.4	2492	1	A44213	nonstructural poly
937	15	48.4	2492	1	C44213	nonstructural poly
938	15	48.4	2492	1	MNWVTD	alpha-51D immobill
939	15	48.4	2533	2	T28675	alpha-51D immobill
940	15	48.4	2533	2	T28674	alpha-51D immobill
941	15	48.4	2543	2	T31687	sulfate antigen - p
942	15	48.4	2643	2	T29149	hypothetical prote
943	15	48.4	2731	1	VFIHJH	genome polypepten
944	15	48.4	2733	2	S15760	genome polypepten
945	15	48.4	2844	2	S28291	hypothetical prote
946	15	48.4	3005	2	T08841	polypepten - dour
947	15	48.4	3010	1	A45573	genome polypepten
948	15	48.4	3010	1	GNNWCJ	genome polypepten
949	15	48.4	3010	1	GNNVTC	genome polypepten
950	15	48.4	3010	1	GNNVTW	genome polypepten
951	15	48.4	3010	1	S18030	genome polypepten
952	15	48.4	3020	2	A43932	mucin 2 precursor,
953	15	48.4	3071	2	T45584	hypothetical prote
954	15	48.4	3414	1	GNNVNE	genome polypepten
955	15	48.4	3507	2	T34513	hypothetical prote
956	15	48.4	3600	2	D86161	Fl1003.12 protein -
957	15	48.4	4116	2	T13719	calo protein - fru
958	15	48.4	4307	2	T20721	hypothetical prote
959	15	48.4	4660	2	T42737	gp130 protein prec
960	15	48.4	5138	2	B96695	hypothetical prote
961	15	48.4	5376	2	T44215	zonadhesin - mouse
962	15	48.4	6658	2	T13931	projecfin - fruit
963	15	48.4	13288	2	T03099	mucin, submaxillar
964	15	48.4	13	2	B58533	CD61 homolog - cha
965	15	48.4	16	2	C59045	alpha-conotoxin Au
966	15	48.4	16	2	A59045	alpha-conotoxin Au
967	15	48.4	22	2	B55538	uvrC protein - Pse
968	15	48.4	24	2	A27128	sperm histone - ra
969	15	48.4	26	1	SMNC	metallochionein -
970	15	48.4	26	1	S55029	CAP3 protein - ant
971	15	48.4	27	2	S55030	CAP3 protein - ant
972	15	48.4	34	2	B81213	hypothetical prote
973	15	48.4	37	2	B31252	metallochionein II
974	15	48.4	41	2	A59149	sigma-conotoxin GV
975	15	48.4	45	1	VSWTAL	purthionin A-1 -
976	15	48.4	46	1	VTVAA3	viscotoxin A3 - Eu
977	15	48.4	46	1	LORDAL	ligetoxin A - Arge
978	15	48.4	47	1	A24074	pyrularia thionin

979 14 45.2 47 2 S22828 protamine - killer
980 14 45.2 47 2 I48949 cellular disintegr
981 14 45.2 49 2 S00228 protamine - horse
982 14 45.2 49 2 S02007 protamine I - rabb
983 14 45.2 50 1 HSPG sperm histone - pi
984 14 45.2 50 1 HSRH sperm histone - sh
985 14 45.2 50 2 S22582 protamine I - sagu
986 14 45.2 50 2 S21672 protamine I - pig
987 14 45.2 50 2 T38209 probable metalloth
988 14 45.2 51 1 HSBOS sperm histone P1 -
989 14 45.2 51 1 HSMSS1 protamine - mouse
990 14 45.2 51 2 S03997 protamine 1 - rat
991 14 45.2 52 2 S65712 metallochionein 1
992 14 45.2 53 2 S24596 metallochionein -
993 14 45.2 53 2 S75905 hypothetical prote
994 14 45.2 55 1 B7AG58 vifb7 protein prec
995 14 45.2 56 1 N1BSSA subfilin precursor
996 14 45.2 56 2 T12783 subclatin 168 prec
997 14 45.2 57 2 S12957 venom animal Kunlt
998 14 45.2 58 1 SMD1S metallochionein 1
999 14 45.2 58 2 G97810 hypothetical prote
1000 14 45.2 59 1 FEDV1V ferredoxin [4Fe-4S]

ALIGNMENTS

RESULT 1

T28945 hypothetical protein F07C4.2 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C/Accession: T28945

R/Miller, N.; Steellies, L.

submitted to the EMBL Data Library, January 1997

A/Description: The sequence of C. elegans cosmid F07C4.

A/Reference number: Z20546

A/Accession: T28945

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1-155 <MIL>

A/Cross-references: UNIPROT:P91214; UNIPARC:UPI000007DB2D; EMBL:U80023; P1DN:AA048012.1;

A/Experimental source: strain Bristol N2; clone F07C4

C/Genetics:

A/Genes: CESP:F07C4.2

A/Map position: 5

A/Intons: 116/1

C/Superfamily: tetranectin; C-type lectin homology

Query Match 58.1%; Score 18; DB 2; Length 155;
Best Local Similarity 22.2%; Pred. No. 17;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 92 CTTASSSSC 100

RESULT 2

T45962 hypothetical protein F7J8.200 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

C/Accession: T45962

R/Bevan, M.; Zimmermann, W.; Gruenisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Lem

submitted to the Protein Sequence Database, January 2000

A/Reference number: Z23018

A/Accession: T45962

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-470 <BBV>

A/Cross-references: UNIPROT:Q9LFB4; UNIPARC:UPI00000A7D3B; EMBL:AL137189

A/Experimental source: cultivar Columbia; BAC clone F7J8

C/Genetics:
A/Map position: 5
A/Intons: 116/2; 138/3; 162/2; 201/3; 226/2; 250/1; 275/1; 288/2; 307/2; 379/3
A/Note: F7J8.200

Query Match 58.1%; Score 18; DB 2; Length 470;
Best Local Similarity 22.2%; Pred. No. 23;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 23 CSSSSATSC 31

RESULT 3

A34598 ecdysone-induced protein E75A - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster

C/Date: 29-Jun-1990 #sequence_revision 29-Jun-1990 #text_change 31-Dec-2004

C/Accession: A34598

R/Segraves, W.A.; Hogness, D.S.

Genes Dev. 4, 204-219, 1990

A/Title: The E75 ecdysone-inducible gene responsible for the 75B early puff in Drosophila

A/Reference number: A34598; PMID:90249727; PMID:2110921

A/Accession: A34598

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1337 <SEG>

A/Cross-references: UNIPROT:P17671; UNIPARC:UPI0000129B00; GB:X51548; NID:G7910; P1DN:CA

C/Genetics:

A/Genes: FlyBase:Rip75B

A/Cross-references: FlyBase:FBgn0000568

C/Keywords: alternative splicing; DNA binding; nucleus; transcription regulation; zinc f

F;243-520/Domain: erba transforming protein homology <ERBA>

Query Match 58.1%; Score 18; DB 2; Length 1237;
Best Local Similarity 22.2%; Pred. No. 30;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 1070 CSSSSASSC 1078

RESULT 4

B34598 ecdysone-induced protein E75B - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster

C/Date: 29-Jun-1990 #sequence_revision 29-Jun-1990 #text_change 31-Dec-2004

C/Accession: B34598

R/Segraves, W.A.; Hogness, D.S.

Genes Dev. 4, 204-219, 1990

A/Title: The E75 ecdysone-inducible gene responsible for the 75B early puff in Drosophila

A/Reference number: A34598; PMID:90249727; PMID:2110921

A/Accession: B34598

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1394 <SEG>

A/Cross-references: UNIPROT:P17672; UNIPARC:UPI0000129B01; GB:X51549; NID:G7912; P1DN:CA

C/Genetics:

A/Genes: FlyBase:Rip75B

A/Cross-references: FlyBase:FBgn0000568

C/Keywords: alternative splicing; DNA binding; nucleus; transcription regulation; zinc

F;401-677/Domain: erba transforming protein homology <ERBA>

Query Match 58.1%; Score 18; DB 2; Length 1394;
Best Local Similarity 22.2%; Pred. No. 31;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 1227 CSSSSASSC 1235

RESULT 5

S05979
steroid hormone receptor homolog - fruit fly (Drosophila melanogaster)
N/Alternate names: puff 75B protein
C/Species: Drosophila melanogaster
C/Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 05-Oct-2004
C/Accession: S05979
R/Feigl, G.; Gram, M.; Pongs, O.
Nucleic Acids Res. 17, 7167-7178, 1989
A/Title: A member of the steroid hormone receptor gene family is expressed in the 20-OH-
A/Reference number: S05979; MUID:90016778; PMID:2508058
A/Accession: S05979
A/Molecule type: mRNA
A/Residues: 1-1443 <PEP>
A/Cross-references: UNIPROT:P13055; UNIPARC:UP10000129B02; EMBL:X15586; NID:97516; PIDN:
C/Genetics:
A/Genes: FlyBase:Fly75B
A/Cross-references: FlyBase:FlyBase:FBgn0000568
A/Map position: 3 75B
C/Keywords: alternative splicing; DNA binding; steroid hormone receptor; transcription

Query Match 58.1%; Score 18; DB 2; Length 1443;
Best Local Similarity 22.2%; Pred. No. 32;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 1276 CXXXXXXC 1284

RESULT 6

T33355
hypothetical protein F16G10.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C/Accession: T33355
R/Gatung, S.; Scheet, P.
submitted to the EMBL Data Library, July 1998
A/Description: The sequence of C. elegans cosmid F16G10.
A/Reference number: Z21329
A/Accession: T33355
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-199 <GAT>
A/Cross-references: UNIPROT:O76588; UNIPARC:UP1000007891D; EMBL:AF077537; PIDN:AAC26272.
A/Experimental source: strain Bristol N2; clone F16G10
C/Genetics:
A/Genes: CRSP:F16G10.4
A/Map position: 2
A/Functions: 56/3; 69/1; 89/1; 133/3
C/Superfamily: Caenorhabditis elegans hypothetical protein F16G10.4

Query Match 54.8%; Score 17; DB 2; Length 199;
Best Local Similarity 22.2%; Pred. No. 51;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 72 CEATSTATC 80

RESULT 7

S63342
hypothetical protein YNR014w - yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein N2060
C/Species: Saccharomyces cerevisiae
C/Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 09-Jul-2004
C/Accession: S63342; S63344; S63340
R/Hauser, C.T.C.; Urbanus, J.H.M.; Planta, R.J.
submitted to the Protein Sequence Database, April 1996
A/Reference number: S63346

A/Accession: S63342
A/Molecule type: DNA
A/Residues: 1-212 <MAU>

A/Cross-references: UNIPROT:P53719; UNIPARC:UP1000013BAB3; EMBL:Z71629; NID:G1302493; P
A/Experimental source: strain S288C
R/Duesterhoeft, A.; Floeth, M.; Fritz, C.; Heues-Neitzel, D.; Hilbert, H.; Moestl, D.
submitted to the Protein Sequence Database, April 1996
A/Reference number: S62944
A/Accession: S63344
A/Molecule type: DNA
A/Residues: 1-212 <DUE>
A/Cross-references: UNIPARC:UP1000013BAB3; EMBL:Z71629; NID:G1302493; PID:e239786; PID:
A/Experimental source: strain S288C
R/Aert, R.; Verhaeselt, P.; Voet, M.; Volckaert, G.
submitted to the Protein Sequence Database, April 1996
A/Reference number: S62910
A/Accession: S63340
A/Molecule type: DNA
A/Residues: 1-26 <AER>
A/Cross-references: UNIPARC:UP10000179C25; EMBL:Z71629; MIPS:YNR014w
C/Genetics:
A/Cross-references: SGD:S0005297
A/Map position: 14R

Query Match 54.8%; Score 17; DB 2; Length 212;
Best Local Similarity 22.2%; Pred. No. 52;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 40 CXXXXXXC 48

RESULT 8

C70812
probable lpqQ protein - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: C70812
R/Cole, S.T.; Brosch, R.; Parthill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Ruter, S.; Steger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: C70812
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-214 <COL>
A/Cross-references: UNIPROT:O53846; UNIPARC:UP10000165256; GB:AL022004; GB:AL123456; NIT
A/Experimental source: strain H37RV
C/Genetics:
A/Genes: lpqQ

Query Match 54.8%; Score 17; DB 2; Length 214;
Best Local Similarity 22.2%; Pred. No. 52;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 15 CAIAITTAC 23

RESULT 9

H72532
hypothetical protein ABE2236 - Aeropyrum pernix (strain K1)
C/Species: Aeropyrum pernix
C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C/Accession: H72532
R/Kawabuchi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai
awa, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999
A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: H72532
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-244 <KAW>
A:Cross-references: UNIPROT:Q9Y9Q2; UNIPARC:UPI000005E27C; DDBJ:AP000063; NID:95105654;
A:Experimental source: strain KI
C:Genetics:
A:Gene: APE2236
C:Superfamily: Aeropyrum pernix hypothetical protein APE2236

Query Match 54.8%; Score 17; DB 2; Length 244;
Best Local Similarity 22.2%; Pred. No. 54;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 107 CSSSSAALC 115

RESULT 10
T40415
hypothetical protein SPBC4_01 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T40415
R:McDougall, R.C.; Rajadream, M.A.; Barrel, B.G.; Brown, S.; Harris, D.
submitted to the EMBL Data Library, October 1999
A:Reference number: Z21927
A:Accession: T40415
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-248 <MCD>
A:Cross-references: UNIPROT:Q9USS9; UNIPARC:UPI000006AF49; EMBL:AL121863; PIDN:CAB58400;
A:Experimental source: strain 972h-; cosmid c4
C:Genetics:
A:Gene: SPDB:SPBC4_01
A:Map position: 2
A:introns: 230/2

Query Match 54.8%; Score 17; DB 2; Length 248;
Best Local Similarity 22.2%; Pred. No. 54;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 146 CLTTSFAC 154

RESULT 11
H71353
conserved hypothetical protein TP0182 - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: H71353
R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uteerback, T.; MCD
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A>Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98352770; PMID:9665876
A:Accession: H71353
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-259 <COL>
A:Cross-references: UNIPROT:Q83212; UNIPARC:UPI000000CA51; GB:AE001202; GB:AE000520; NID
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0182
C:Superfamily: RNA-binding protein, Yrdc type

Query Match 54.8%; Score 17; DB 2; Length 259;
Best Local Similarity 22.2%; Pred. No. 55;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 31 CASASLSSC 39

RESULT 12
AC0661
probable lipoprotein SRY1394 [imported] - Salmonella enterica subsp. enterica serovar Ty
C:Species: Salmonella enterica subsp. enterica serovar Typh
A>Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AC0661
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.,
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AC0661
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-281 <PAR>
A:Cross-references: UNIPARC:UPI0000059FOE; GB:AL513382; PIDN:CAD01660.1; PID:g16502512;
C:Genetics:
A:Gene: SRY1394

Query Match 54.8%; Score 17; DB 2; Length 281;
Best Local Similarity 22.2%; Pred. No. 56;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 208 CATAAABC 216

RESULT 13
T48975
xyloglucan endo-transglucosylase - Arabidopsis thaliana
N:Alternate names: protein F14D17.60
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: T48975
R:Jordan, N.; Bangert, S.; Wiedemann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Rudd, S.;
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z25008
A:Accession: T48975
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-293 <JOR>
A:Cross-references: UNIPROT:P93046; UNIPARC:UPI00000487CB; EMBL:AL353992; GSPDB:GN00061
A:Experimental source: cultivar Columbia; BAC clone F14D17
C:Genetics:
A:Gene: ATSP:F14D17.60
A:Map position: 3
A:introns: 68/1; 101/3; 174/2
C:Superfamily: endoxylglucan transferase

Query Match 54.8%; Score 17; DB 2; Length 293;
Best Local Similarity 22.2%; Pred. No. 57;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 238 CTRDSSSC 246

RESULT 14
T01948

hypothetical protein F1104.13 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004
C/Accession: T01948
R:/Abu-Threiden, J.; Stoneking, T.; Langston, Y.; Trevasakis, E.
submitted to the EMBL Data Library, October 1998
A/Description: The sequence of A. thaliana F1104.
A/Reference number: Z14466
A/Accession: T01948
A/Status: translated from GB/EMBL/DDBT
A/Molecule type: DNA
A/Residues: 1-298 <ABU>
A/Cross-references: UNIPROT:O82599; UNIPARC:UPI00000A5CDC; EMBL:AF096370; NID:g3695372;
C/Genetics:
A/Map position: 4
A/Intons: 158/3
A/Note: F1104.13
C/Superfamily: Arabidopsis thaliana hypothetical protein F21F14.120

Query Match 54.8%; Score 17; DB 2; Length 298;
Best Local Similarity 22.2%; Pred. No. 57;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 200 CXXXXXXC 208

RESULT 15
A85019
probable bHLH DNA-binding protein [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C/Accession: A85019
R:/Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A/Reference number: A85001; MUID:20083488; PMID:10617198
A/Accession: A85019
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-315 <STO>
A/Cross-references: UNIPROT:Q9M128; UNIPARC:UPI000009FFA; GB:NC_001268; NID:g7268189; F
C/Genetics:
A/Gene: Atg01460
A/Map position: 4
C/Superfamily: Arabidopsis thaliana hypothetical protein F21F14.120

Query Match 54.8%; Score 17; DB 2; Length 315;
Best Local Similarity 22.2%; Pred. No. 58;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 188 CXXXXXXC 196

RESULT 16
AB1224
cobalamin biosynthesis protein Cbid homolog cbid [imported] - Listeria monocytogenes (st
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AB1224
R:/Glaeser, P.; Frangul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feihl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Krefl, U.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltournam, A.; M
Ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AB1224

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-373 <GLA>
A/Cross-references: UNIPROT:Q9Y757; UNIPARC:UPI0000054CCA; GB:NC_003210; PIDN:CAC99272.
C/Genetics:
A/Genetic source: strain EGD-e
C/Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 54.8%; Score 17; DB 2; Length 373;
Best Local Similarity 22.2%; Pred. No. 61;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 21 CAAAAKAC 29

RESULT 17
AD1577
cobalamin biosynthesis protein Cbid homolog cbid [imported] - Listeria innocua (strain
C/Species: Listeria innocua
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AD1577
R:/Glaeser, P.; Frangul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feihl, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Krefl, U.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltournam, A.; M
Ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AD1577
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-373 <GLA>
A/Cross-references: UNIPROT:Q92C44; UNIPARC:UPI00000CC44; GB:AL592022; PIDN:CAC93688.1
C/Genetics:
A/Gene: cbid
C/Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 54.8%; Score 17; DB 2; Length 373;
Best Local Similarity 22.2%; Pred. No. 61;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 21 CAAAAKAC 29

RESULT 18
T43352
nuclear receptor NHR-13 - Caenorhabditis elegans (fragment)
C/Species: Caenorhabditis elegans
C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C/Accession: T43352
R:/Sluder, A.E.; Mathews, S.W.; Hough, D.; Yin, V.P.; Maina, C.V.
Genome Res. 9, 103-120, 1999
A/Title: The nuclear receptor superfamily has undergone extensive proliferation and div
A/Reference number: Z22443; MUID:99148134; PMID:10022975
A/Accession: T43352
A/Status: preliminary; translated from GB/EMBL/DDBT
A/Molecule type: mRNA
A/Residues: 1-415 <SLU>
A/Cross-references: UNIPROT:Q9XYB8; UNIPARC:UPI0000075586; EMBL:AF083228; NID:g4139081;
C/Genetics:
A/Gene: nhr-13
A/Map position: 5

Query Match 54.8%; Score 17; DB 2; Length 415;
Best Local Similarity 22.2%; Pred. No. 63;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 12 CSSSSNSSC 20

RESULT 19

T33024 hypothetical protein K07H8.2 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C:Accession: T33024

R:Fulton, B.; Hawkins, J.; Gattung, S.; Woldmann, P.; Elliott, G.

submitted to the EMBL Data Library, February 1998

A:Description: The sequence of *C. elegans* cosmid K07H8.

A:Reference number: Z21264

A:Accession: T33024

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-507 <RUF>

A:Cross-references: UNIPROT:Q45182; UNIPARC:UPI0000080E7E; EMBL:AF047659; PIDD:AA04427.

A:Experimental source: strain Bristol N2; clone K07H8

C:Genetics:

A:Gene: CESP:K07H8.2

A:Map position: 4

A:Introns: 20/3; 55/2; 93/3; 152/3; 447/3

Query Match 54.8%; Score 17; DB 2; Length 507;
Best Local Similarity 22.2%; Pred. No. 66;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 210 CASSIATAC 218

RESULT 20

T48004 multifunctional aminoacyl-tRNA ligase-like protein - *Arabidopsis thaliana*

N:Alternate names: protein T17J13.80

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004

C:Accession: T48004

R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Lemcke, K.; Mayer,

submitted to the Protein Sequence Database, February 2000

A:Reference number: Z24482

A:Accession: T48004

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-530 <RIF>

A:Cross-references: UNIPROT:Q9M1R2; UNIPARC:UPI000009CD43; EMBL:AL138651

A:Experimental source: cultivar Columbia; BAC clone T17J13

C:Genetics:

A:Map position: 3

A:Introns: 64/3; 94/3; 133/3; 196/2; 218/3; 259/3; 310/3; 405/3; 478/3; 510/1

A>Note: T17J13.80

C:Superfamily: proline-tRNA ligase pros

Query Match 54.8%; Score 17; DB 2; Length 530;
Best Local Similarity 22.2%; Pred. No. 67;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 354 CTATGATAC 362

RESULT 21

A42143 skn-1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: A42143; T29535

R:Bowerman, B.; Eaton, B.A.; Priess, J.R.

Cell 68, 1061-1075, 1992

A>Title: skn-1, a maternally expressed gene required to specify the fate of ventral blas

A:Reference number: A42143; MUID:92191285; PMID:1547503

A:Accession: A42143

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-533 <BOW>

A:Cross-references: UNIPROT:Q8MPW3; UNIPARC:UPI0000081925

A>Note: sequence extracted from NCBI backbone (NCBIN:88973, NCBI:P:88974)

R:Pauley, A.; Gattung, S.

submitted to the EMBL Data Library, December 1995

A:Description: The sequence of *C. elegans* cosmid T19E7.

A:Reference number: Z20637

A:Accession: T29535

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-386 <PAU>

A:Cross-references: UNIPARC:UPI000017A521; EMBL:U42843; PIDD:AAA83594.1

C:Genetics:

A:Gene: CESP:skn-1

A:Introns: 141/3; 249/1; 300/1

C:Keywords: nucleus

F:495-532/Domain: fos/jun DNA-binding domain homology <RUD>

Query Match 54.8%; Score 17; DB 2; Length 533;
Best Local Similarity 22.2%; Pred. No. 67;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 384 CTTDSSTC 392

RESULT 22

T07739

probable ferredoxinase (EC 4.99.1.1) - potato

C:Species: *Solanum tuberosum* (potato)

C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004

C:Accession: T07739

R:Johnston, D.J.

submitted to the EMBL Data Library, April 1998

A:Reference number: Z15932

A:Accession: T07739

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-543 <JOH>

A:Cross-references: UNIPROT:Q64391; UNIPARC:UPI00000A543B; EMBL:AJ005802; PIDD:CAA06705

A:Experimental source: cv. Bintje

C:Function:

A:Description: catalyzes the insertion of iron into protoporphyrin to produce heme

A:Pathway: heme biosynthesis

C:Superfamily: ferredoxinase

C:Keywords: iron; lyase; mitochondrial inner membrane

Query Match 54.8%; Score 17; DB 2; Length 543;
Best Local Similarity 22.2%; Pred. No. 67;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 31 CTSSFAASC 39

RESULT 23

THHUB

thrombomodulin precursor [validated] - human

C:Species: *Homo sapiens* (man)

C>Date: 31-Dec-1988 #sequence_revision 12-May-1995 #text_change 09-Jul-2004

C:Accession: A41442; A28307; A29680; A27073; JX0264; S3554

R:Shitani, T.; Shiojiri, S.; Ito, H.; Yamamoto, S.; Kusumoto, H.; Deyashiki, Y.; Maruyama,

J. Biochem. 103, 281-285, 1988

A>Title: Gene structure of human thrombomodulin, a cofactor for thrombin-catalyzed acti-

A:Reference number: A41442; MUID:88227901; PMID:2836377
 A:Accession: A41442
 A:Molecule type: DNA
 A:Residues: 1-575 <SHI>
 A:Cross-references: UNIPROT:P07204; UNIPARC:UPI00000498EB; DDBJ:D00210; NID:g220126; PID:R:Jackman, R.W.; Beeler, D.L.; Filtze, L.; Soff, G.; Rosenberg, R.D.
 Proc. Natl. Acad. Sci. U.S.A. 84, 6425-6429, 1987
 A:Title: Human thrombomodulin gene is intron depleted; nucleic acid sequences of the cDN
 A:Reference number: A28307; MUID:87317665; PMID:2819876
 A:Accession: A28307
 A:Molecule type: DNA; mRNA
 A:Residues: 1-472, 'A', 474-575 <JAC>
 A:Cross-references: UNIPARC:UPI00000002BD; GB:J02973; NID:g339658; PIDN:AAA61175.1; PID:R:Suzuki, K.; Kusumoto, H.; Deyshehik, Y.; Nishioka, J.; Matsuyama, I.; Zushi, M.; Kawaba
 EMBO J. 6, 1891-1897, 1987
 A:Title: Structure and expression of human thrombomodulin, a thrombin receptor on endoth
 A:Reference number: A29680; MUID:88004395; PMID:2820710
 A:Accession: A29680
 A:Molecule type: mRNA
 A:Residues: 1-575 <SUZ>
 A:Cross-references: UNIPARC:UPI00000498PB; GB:X05495; NID:g37123; PIDN:CAA29045.1; PID:G
 A:Experimental source: lung endothelium
 A:Note: part of this sequence, including the amino end of the mature protein, were deter
 R:Wen, D.; Dittman, W.A.; Ye, R.D.; Deaven, L.L.; Majerus, P.W.; Sadler, J.E.
 Biochemistry 26, 4350-4357, 1987
 A:Title: Human thrombomodulin: complete cDNA sequence and chromosome localization of the
 A:Reference number: A27073; MUID:88024950; PMID:2822087
 A:Accession: A27073
 A:Molecule type: mRNA
 A:Residues: 1-472, 'A', 474-575 <WEN>
 A:Cross-references: UNIPARC:UPI00000002BD; GB:M16552; NID:g339656; PIDN:AAB59508.1; PID:G
 A:Experimental source: placenta
 A:Note: parts of this sequence were determined by protein sequencing
 R:Tamamoto, S.; Mizoguchi, T.; Tamaki, T.; Ohkuchi, M.; Kimura, S.; Aoki, N.
 J. Biochem. 113, 433-440, 1993
 A:Title: Urinary thrombomodulin, its isolation and characterization.
 A:Reference number: JX0264; MUID:93293792; PMID:8390446
 A:Accession: JX0264
 A:Molecule type: protein; mRNA
 A:Residues: 19-472, 'A', 474-486 <YAM>
 A:Cross-references: UNIPARC:UPI000017341
 A:Experimental source: urine
 A:Note: the urinary form appears to be identical with that circulating in plasma
 R:Gerlitz, B.; Hassell, T.; Vlahos, C.J.; Parkinson, J.F.; Bang, N.U.; Grinnell, B.W.
 Biochem. J. 295, 131-140, 1993
 A:Title: Identification of the predominant glycosaminoglycan-attachment site in soluble
 A:Reference number: S38954; MUID:94029900; PMID:8216207
 A:Accession: S38954
 A:Molecule type: protein
 A:Residues: 475-491, 'X', 493-494 <GER>
 A:Cross-references: UNIPARC:UPI000017342
 A:Note: the residue designated 'X' was determined to be a Ser with covalently bound chon
 R:Meininger, D.P.; Komives, E.A.
 submitted to the Brookhaven Protein Data Bank, September 1995
 A:Reference number: A67369; PDB:1ZAO
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R:Fullinsky, A.; Mathews, I.I.
 submitted to the Brookhaven Protein Data Bank, August 1994
 A:Reference number: A52804; PDB:1HLT
 A:Contents: annotation; X-ray crystallography, 3.0 angstroms, residues 426-442
 R:Hrabal, R.; Komives, E.A.; Ni, F.
 submitted to the Brookhaven Protein Data Bank, November 1995
 A:Reference number: A65583; PDB:1FGD
 A:Contents: annotation; conformation by (1)H-NMR, residues 427-444
 R:Hrabal, R.; Komives, E.A.; Ni, F.
 Protein Sci. 5, 195-203, 1996
 A:Title: Structural resiliency of an EGF-like subdomain bound to its target protein, the
 A:Reference number: A58595; MUID:96276211; PMID:8745396
 A:Contents: annotation; conformation by (1)H-NMR
 C:Genetics:
 A:Gene: GDB:THBD
 A:Cross-references: GDB:119613; OMIM:188040

A:Map position: 20p11.2-20p11.2
 A:introns: #status absent
 C:Complex: homodimer, urinary form
 C:Function:
 A:Description: inhibits thrombin activation of fibrinogen; cofactor for thrombin activa
 A:Pathway: blood coagulation/moderation
 A:Note: the membrane-bound form is located on the endothelium luminal surface of arteri
 A:Note: thrombin complexed with the membrane-bound form is subject to endocytosis
 C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology
 C:Keywords: anticoagulant; beta-hydroxyaspartate; beta-hydroxyaspartic acid; blood coag
 e protein
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-575/Product: thrombomodulin, membrane-bound form #status predicted <MAT>
 F:19-513/Domain: extracellular #status predicted <EXT>
 F:19-486/Product: thrombomodulin, urinary form #status experimental <MAU>
 F:24-157/Domain: C-type lectin homology <LCH>
 F:177-199/Region: PEST sequence
 F:201-233/Region: PEST sequence
 F:245-280/Domain: EGF homology <EG1>
 F:288-323/Domain: EGF homology <EG2>
 F:329-362/Domain: EGF homology <EG3>
 F:369-404/Domain: EGF homology <EG4>
 F:408-439/Domain: EGF homology <EG5>
 F:445-480/Domain: EGF homology <EG6>
 F:485-513/Region: PEST sequence
 F:517-539/Domain: transmembrane #status predicted <TMN>
 F:540-575/Domain: intracellular #status predicted <INT>
 F:47,115,116,382,409/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:174,125,411,504/Binding site: carbohydrate (Thr) (covalent) #status predicted
 F:245-256,252-265,267-280,288-296,292-308,310-323,329-340,336-349,351-362,369-378,374-38
 F:334,498/Binding site: carbohydrate (Ser) (covalent) #status predicted
 F:342/Modified site: erythro-beta-hydroxyaspartate (Asn) #status experimental
 F:490,492/Binding site: chondroitin sulfate (Ser) (covalent) (partial) #status experimen
 Query Match 54.8%; Score 17; DB 1; Length 575;
 Best Local Similarity 22.2%; Pred. No. 69;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 12
 DB 280 CTASATQSC 288
 RESULT 24
 T07953
 lectin-like protein zsp2, zygote-specific - Chlamydomonas reinhardtii
 C:Species: Chlamydomonas reinhardtii
 C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004
 C:Accession: T07953
 R:Suzuki, L.; Woessner, J.P.; Uchida, H.; Kuriyama, H.; Takano, H.; Kuroiwa, H.; Yuasa,
 submitted to the EMBL Data Library, March 1998
 A:Description: A lectin-like protein mediates the assembly of the extracellular matrix c
 A:Reference number: Z16243
 A:Accession: T07953
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-582 <SUZ>
 A:Cross-references: UNIPROT:O65153; UNIPARC:UPI00000454CF; EMBL:AF053099; NID:g2997677;
 C:Genetics:
 A:Gene: zsp2
 A:introns: 27/1; 292/3
 Query Match 54.8%; Score 17; DB 2; Length 582;
 Best Local Similarity 22.2%; Pred. No. 69;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 4 CXXXXXXC 12
 DB 257 CARSTTTC 265
 RESULT 25
 T07952

lectin-like protein, zygote-specific - Chlamydomonas reinhardtii
 C:Species: Chlamydomonas reinhardtii
 C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004
 C:Accession: T07952
 R:Stuiki, L.; Weesener, J.P.; Uchida, H.; Kuriyama, H.; Takano, H.; Kuroiwa, H.; Yuasa, R.
 Submitted to the EMBL Data Library, March 1998
 A:Description: A lectin-like protein mediates the assembly of the extracellular matrix
 A:Reference number: Z16243
 A:Accession: T07952
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-582 <SUZ>
 A:Cross-references: UNIPROT:Q9SB11; UNIPARC:UPI000004454C; EMBL:AF053098; NID:g2997675;
 C:Genetics:
 A:Gene: zsp2
 A:Introns: 27/1; 292/3

Query Match 54.8%; Score 17; DB 2; Length 582;
 Best Local Similarity 22.2%; Pred. No. 69;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12
 Db 257 CARSTTTC 265

RESULT 26
 B25682
 homeotic protein engrailed - fruit fly (Drosophila virilis)
 C:Species: Drosophila virilis
 C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004
 C:Accession: B25682
 R:Kasib, J.A.; Poole, S.J.; Wright, D.K.; O'Farrell, P.H.
 EMBO J. 5, 3583-3589, 1986
 A:Title: Sequence conservation in the protein coding and intron regions of the engrailed
 A:Reference number: A91059; MUID:87161768; PMID:2881781
 A:Accession: B25682
 A:Molecule type: DNA
 A:Residues: 1-584 <KAS>
 A:Cross-references: UNIPROT:P09145; UNIPARC:UPI000012CA14; GB:X04727; NID:g9173; PIDN:CA
 C:Genetics:
 A:Gene: en
 A:Cross-references: FlyBase:FBgn013111
 A:Introns: 470/1; 502/3
 C:Superfamily: engrailed homeotic protein; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:487-543/Domain: homeobox homology <HOX>

Query Match 54.8%; Score 17; DB 2; Length 584;
 Best Local Similarity 22.2%; Pred. No. 69;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12
 Db 384 GSASSSSGC 392

RESULT 27
 S41011
 hypothetical protein ZK757.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
 C:Accession: S41011
 R:Thomas, K.
 submitted to the EMBL Data Library, December 1993
 A:Reference number: S41011
 A:Accession: S41011
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-589 <THO>
 A:Cross-references: UNIPROT:P34679; UNIPARC:UPI000013BC25; EMBL:Z29121; NID:g438366; PID
 C:Genetics:
 A:Introns: 19/2; 54/3; 114/3; 146/3; 178/3; 209/1; 253/1; 302/3; 324/3; 352/1; 392/3; 48

C:Keywords: transmembrane protein

Query Match 54.8%; Score 17; DB 2; Length 589;
 Best Local Similarity 22.2%; Pred. No. 69;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12
 Db 100 CITSSSTTC 108

RESULT 28
 A42086
 CD30 antigen precursor - human
 N:Alternate names: Ki-1 antigen; nerve growth factor receptor family member CD30
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
 C:Accession: A42086
 R:Durkop, H.; Latza, U.; Hummel, M.; Eitelbach, F.; Seed, B.; Stein, H.
 Cell 68, 421-427, 1992
 A:Title: Molecular cloning and expression of a new member of the nerve growth factor rec
 A:Reference number: A42086; MUID:92154659; PMID:1310894
 A:Accession: A42086
 A:Molecule type: mRNA
 A:Residues: 1-595 <DUR>
 A:Cross-references: UNIPROT:P28908; UNIPARC:UPI0000000971; GB:M83554; NID:g180095; PIDN:
 A:Experimental source: HUT-102 cell line
 A:Note: sequence extracted from NCBI backbone (NCBIN:82088, NCBI:P:82090)
 C:Genetics:
 A:Gene: GDB:CD30; DIS166E
 A:Cross-references: GDB:131547; OMIM:153243
 A:Map position: 1p36-1p36
 C:Keywords: glycoprotein; growth factor receptor; transmembrane protein
 F:1-18/Domain: signal sequence #status predicted <Sig>
 F:19-383/Domain: extracellular #status predicted <EXT>
 F:384-407/Domain: transmembrane #status predicted <TM>
 F:408-595/Domain: intracellular #status predicted <CYT>
 F:101,276/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 54.8%; Score 17; DB 2; Length 595;
 Best Local Similarity 22.2%; Pred. No. 69;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12
 Db 289 CATSATNSC 297

RESULT 29
 D83286
 hypothetical protein PA2886 [imported] - Pseudomonas aeruginosa (strain PAO1)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C:Accession: D83286
 R:Stoyer, C.K.; Pham, X.O.; Edwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Ba
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, J.N.; Folger, K.R.; Kae, A.; Lartig, K.; Lm,
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathc
 A:Reference number: A83250; MUID:20437337; PMID:10984043
 A:Accession: D83286
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-600 <STO>
 A:Cross-references: UNIPROT:Q9H2M1; UNIPARC:UPI00000C5861; GB:AE004714; GB:AE004091; NIT
 A:Experimental source: strain PAO1
 C:Genetics:
 A:Gene: PA2886

Query Match 54.8%; Score 17; DB 2; Length 600;
 Best Local Similarity 22.2%; Pred. No. 69;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 92 CAAALAAC 100

RESULT 30

T15408
hypothetical protein C04F6.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T15408

R:Nhan, M.
submitted to the EMBL Data Library, December 1995
A:Description: The sequence of C. elegans cosmid C04F6.
A:Reference number: Z18346

A:Accession: T15408
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-617 <NHA>
A:Cross-references: UNIPROT:Q11174; UNIPARC:UPI0000060EB1; EMBL:U42835; NID:G125760; P
C:Genetics:
A:Gene: CESP.C04F6.3
A:Introns: 28/1; 66/2; 504/1

Query Match 54.8%; Score 17; DB 2; Length 617;
Best Local Similarity 22.2%; Pred. No. 70;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 524 CTTSTASGC 532

RESULT 31

A36962
laccase (EC 1.10.3.2) precursor - fungus (*Filobasidium floricorne*) (ATCC 34873)
N:Alternate names: diphenol oxidase

C:Species: *Filobasidium floricorne*, *Cryptococcus neoformans*
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 31-Dec-2004
C:Accession: A36962

R:Williamson, P.R.
J. Bacteriol. 176, 656-664, 1994
A:Title: Biochemical and molecular characterization of the diphenol oxidase of *Cryptococcus*
A:Reference number: A36962; MUID:94131944; PMID:8300520
A:Accession: A36962

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-624 <MIL>
A:Cross-references: UNIPROT:Q9UQZ7; UNIPARC:UPI00001751D2; GB:I22866
C:Genetics:
A:Gene: CNLAC1

A:Introns: 25/3; 69/3; 117/1; 137/3; 176/2; 204/2; 223/3; 238/3; 371/2; 391/2; 426/3; 55
C:Keywords: copper; glycoprotein; oxidoreductase

Query Match 54.8%; Score 17; DB 2; Length 624;
Best Local Similarity 22.2%; Pred. No. 70;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 243 CTATGSSSC 251

RESULT 32

A41029
integrin beta-8 chain precursor - human

C:Species: *Homo sapiens* (man)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 31-Dec-2004
C:Accession: A41029

R:Moyle, M.; Napier, M.A.; McLean, J.W.
J. Biol. Chem. 266, 19650-19658, 1991
A:Title: Cloning and expression of a divergent integrin subunit beta-8.
A:Reference number: A41029; MUID:92011767; PMID:1918072

A:Accession: A41029
A:Molecule type: mRNA
A:Residues: 1-769 <MOY>

A:Cross-references: UNIPROT:P26012; UNIPARC:UPI000012DA14; GB:M73780; NID:G184520; PID:
C:Superfamily: Integrin, beta subunit, laminin-type BGF-like homology
C:Keywords: cell adhesion; cytoskeleton; duplication; extracellular matrix; glycoprotein

Query Match 54.8%; Score 17; DB 2; Length 769;
Best Local Similarity 22.2%; Pred. No. 74;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 47 CASSNASC 55

RESULT 33

A46140
diacylglycerol kinase (EC 2.7.1.107) - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*

C:Date: 22-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A46140

R:Masai, I.; Hosoya, T.; Kojima, S.; Hotta, Y.
Proc. Natl. Acad. Sci. U.S.A. 89, 6030-6034, 1992
A:Title: Molecular cloning of a *Drosophila* diacylglycerol kinase gene that is expressed
A:Reference number: A46140; MUID:92335231; PMID:1321433

A:Accession: A46140
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-791 <MAS>
A:Cross-references: UNIPROT:Q01583; UNIPARC:UPI000016B61; GB:D11120; NID:G217332; PID:
A:Note: sequence extracted from NCBI backbone (NCBIN:108569, NCBI:P:108571)
C:Genetics:

A:Gene: FlyBase:Dgk
A:Cross-references: FlyBase:Fgn0004568
C:Superfamily: Fruit fly diacylglycerol kinase; calmodulin repeat homology; protein kin
C:Keywords: EF hand; phosphotransferase

Query Match 54.8%; Score 17; DB 2; Length 791;
Best Local Similarity 22.2%; Pred. No. 75;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 18 CATPAAVAC 26

RESULT 34

B46140
diacylglycerol kinase (EC 2.7.1.107) - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 29-Sep-1999
C:Accession: B46140

R:Masai, I.; Hosoya, T.; Kojima, S.; Hotta, Y.
Proc. Natl. Acad. Sci. U.S.A. 89, 6030-6034, 1992
A:Title: Molecular cloning of a *Drosophila* diacylglycerol kinase gene that is expressed
A:Reference number: A46140; MUID:92335231; PMID:1321433

A:Accession: B46140
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-796 <MAS>
A:Cross-references: UNIPARC:UPI00001756B6; GB:D11120
C:Genetics:

A:Gene: FlyBase:Dgk
A:Cross-references: FlyBase:Fgn0004568
C:Superfamily: Fruit fly diacylglycerol kinase; calmodulin repeat homology; protein kin
C:Keywords: EF hand; phosphotransferase

Query Match 54.8%; Score 17; DB 2; Length 796;
Best Local Similarity 22.2%; Pred. No. 75;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12

Db 18 CATAAAVAC 26

RESULT 35

Integrin beta-5 chain precursor - human

C/Species: Homo sapiens (man)

C/Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 31-Dec-2004

C/Accession: A38308; A35775; S12534; S11708

R/McClean, J.W.; Vestral, D.J.; Chersesh, D.A.; Bodary, S.C.

J Biol Chem. 265, 17126-17131, 1990

A/Title: cDNA sequence of the human integrin beta-5 subunit.

A/Reference number: A38308; M01D:91009141; PMID:2211615

A/Accession: A38308

A/Molecule type: mRNA

A/Residues: 1-799 <MCL>

A/Cross-references: UNIPROT:P18084; UNIPARC:UPI00004D05B; GB:J05633; NID:9186504; PIDN:

A/Note: parts of this sequence, including the amino end of the mature protein, were conf

R/Suzuki, S.; Huang, Z.S.; Tanihara, H.

Proc. Natl. Acad. Sci. U.S.A. 87, 5354-5358, 1990

A/Title: Cloning of an integrin beta subunit exhibiting high homology with integrin beta

A/Reference number: A35775; M01D:90319111; PMID:2371275

A/Accession: A35775

A/Molecule type: mRNA

A/Residues: 1-132, 'A', 194-644, 'L', 646-789, 793-799 <SUZ>

A/Cross-references: UNIPARC:UPI000016AADB; GB:M35011; NID:9184524; PIDN:AAAS2707.1; PID:

R/Ramaswamy, H.; Hemler, M.E.

EMBO J. 9, 1561-1568, 1990

A/Title: Cloning, primary structure and properties of a novel human integrin beta subunit

A/Reference number: S12534; M01D:90228356; PMID:2328726

A/Accession: S12534

A/Molecule type: mRNA

A/Residues: 1-644, 'L', 646-799 <RAM>

A/Cross-references: UNIPARC:UPI000012DA10; EMBL:X53002; NID:933952; PIDN:CAA37188.1; PID:

C/Genetics:

A/Gene: GDB:ITGB5

A/Cross-references: GDB:128005; OMIM:147561

A/Map position: 17q11-17qter

C/Superfamily: Integrin, beta subunit; laminin-type EGF-like homology

C/Keywords: cell adhesion; cytoskeleton; duplication; extracellular matrix; glycoprotein

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-799/Product: integrin beta-5 chain #status experimental <MAT>

F:425-719/Domain: extracellular #status predicted <EXT>

F:463-513/Domain: laminin-type EGF-like homology <LEG>

F:720-742/Domain: transmembrane #status predicted <TM>

F:743-799/Domain: intracellular #status predicted <INT>

F:347,460,477,505,552,586,654,705/Binding site: carbohydrate (Asn) (covalent) #status pr

Query Match 54.8%; Score 17; DB 2; Length 799;

Best Local Similarity 22.2%; Pred. No. 75;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXC 12

Db 28 CTSGSATSC 36

RESULT 36

hypothetical protein FA110.140 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C/Accession: T05201

R/Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Hobeisel, J.; Mew

submitted to the Protein Sequence Database, February 1999

A/Reference number: Z15402

A/Accession: T05201

A/Molecule type: DNA

A/Residues: 1-934 <BEV>

A/Cross-references: UNIPROT:Q9SMY8; UNIPARC:UPI000009D8D1; EMBL:AL035525

A/Experimental source: cultivar Columbia; BAC clone FA110

C/Genetics:

A/Map position: 4

A/Intons: 123/2; 203/3; 248/1; 281/2; 347/2; 437/3; 470/3; 500/2; 533/3; 590/3; 641/3;

A/Note: FA110.140

Query Match 54.8%; Score 17; DB 2; Length 934;

Best Local Similarity 22.2%; Pred. No. 79;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXC 12

Db 701 CTSATTASC 709

RESULT 37

hypothetical protein U30 - human herpesvirus 7 (strain J1)

C/Species: human herpesvirus 7

A/Variety: strain J1

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C/Accession: T41932

R/Nicholas, J.

submitted to the EMBL Data Library, December 1995

A/Description: Determination and analysis of the complete nucleotide sequence of human h

A/Reference number: Z22022

A/Accession: T41932

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-938 <NIC>

A/Cross-references: UNIPROT:P52438; UNIPARC:UPI0000007D8; EMBL:U43400; PIDN:AMC54692.1

A/Experimental source: strain J1

C/Genetics:

A/Note: U30

Query Match 54.8%; Score 17; DB 2; Length 938;

Best Local Similarity 22.2%; Pred. No. 79;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXC 12

Db 689 CSSTATTTC 697

RESULT 38

hypothetical protein D1044.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C/Accession: T15881

R/Pauley, A.

submitted to the EMBL Data Library, June 1994

A/Description: The sequence of C. elegans cosmid D1044.

A/Reference number: Z18423

A/Accession: T15881

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1895 <PAU>

A/Cross-references: UNIPARC:UPI000004CAE8; EMBL:U00065; NID:9495681; PID:9495684; PIDN:J

A/Experimental source: strain Bristol N2

C/Genetics:

A/Gene: CESP:D1044.3

A/Intons: 35/2; 63/2; 150/1; 191/3; 207/2; 405/3; 551/3; 629/1; 764/1; 933/1; 959/2; 10

Query Match 54.8%; Score 17; DB 2; Length 1895;

Best Local Similarity 22.2%; Pred. No. 96;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXC 12

Db 1233 CTSSSSSSC 1241

RESULT 39

S50820


```

surface protein type 51B - Parametium tetraurelia
C:Species: Parametium tetraurelia
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S50820
R:Scott, J.; Leeck, C.; Forney, J.
Nucleic Acids Res. 22, 5079-5084, 1994
A:Title: Analysis of the micronuclear B type surface protein gene in Parametium tetraurelia
A:Reference number: S50820; MUID:95096630; PMID:7800503
A:Accession: S50820
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2395 <SCO>
A:Cross-references: UNIPROT:Q27167; UNIPARC:UPI00007DPAF6; EMBL:U07603; NID:g467226; PID
C:Genetics:
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1994
C:Genetics:
A:Genetic code: SGC5
A:introns: 472/3; 1310/3; 1821/3
C:Superfamily: G surface protein

Query Match          54.8%; Score 17; DB 1; Length 2395;
Best Local Similarity 22.2%; Pred. No. 1e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      4 CXXXXXXXC 12
      |
DB      1408 CASLTSTC 1416

RESULT 40
S09118
G surface protein 168 - Parametium primaurelia
C:Species: Parametium primaurelia
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C:Accession: S09118
R:Prat, A.
J. Mol. Biol. 211, 521-535, 1990
A:Title: Conserved sequences flank variable tandem repeats in two alleles of the G surface
A:Reference number: S09118; MUID:90172419; PMID:2308165
A:Accession: S09118
A:Molecule type: DNA
A:Residues: 1-2704 <PRA>
A:Cross-references: UNIPROT:P17053; UNIPARC:UPI000012AE23; EMBL:X52133; NID:g10049; PIDN
C:Genetics:
A:Genetic code: SGC5
C:Superfamily: G surface protein

Query Match          54.8%; Score 17; DB 2; Length 2704;
Best Local Similarity 22.2%; Pred. No. 1.1e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      4 CXXXXXXXC 12
      |
DB      1106 CSTSTAGTC 1114

RESULT 41
A23475
G surface protein - Parametium primaurelia
C:Species: Parametium primaurelia
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004
C:Accession: A23475
R:Prat, A.; Katinka, M.; Caron, F.; Meyer, E.
J. Mol. Biol. 189, 47-60, 1986
A:Title: Nucleotide sequence of the Parametium primaurelia G surface protein. A huge pro
A:Reference number: A23475; MUID:87060934; PMID:3783679
A:Accession: A23475
A:Molecule type: DNA
A:Residues: 1-2718 <PRA>
A:Cross-references: UNIPROT:P13837; UNIPARC:UPI0000177F7E
A:Note: the authors translated the codon TGC for residue 2665 as Trp
C:Genetics:
A:Genetic code: SGC5
C:Superfamily: G surface protein

```

[illegible]

RESULT 44
H66753
propinase p12 protein 25 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: H66753
R:Boletín, A.; Winkler, P.; Manger, S.; Jallón, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp
A:Reference number: A86625; MWID:21235186; PMID:11337471
A:Accession: H66753
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-87 <STO>
A:Cross-references: UNIPROT:Q9CGR4; UNIPARC:UPI000006978; GB:AE005176; PID:g12723983; F
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: p1225

Query Match 51.6%; Score 16; DB 2; Length 87;
Best Local Similarity 22.2%; Pred. No. 1.1e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 30 CTKASLSC 38

RESULT 45
J50036
Clara cell 10K protein precursor - human
N:Alternate names: urinary protein 1
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: J50036; P50309; A56890; I38397
R:Singh, G.; Katyal, S.L.; Brown, W.E.; Phillips, S.; Kennedy, A.L.; Anthony, J.; Squeglia
Biochim. Biophys. Acta 950, 329-337, 1998
A:Title: Amino-acid and cDNA nucleotide sequences of human Clara cell 10KDa protein.
A:Reference number: J50036; MWID:89000784; PMID:3167058
A:Accession: J50036
A:Molecule type: mRNA
A:Residues: 1-91 <STN>
A:Cross-references: UNIPROT:P11684; UNIPARC:UPI0000036901; GB:X13197; NID:g23131; PID:N.C
A:Accession: P50309
A:Molecule type: protein
A:Residues: 22-23, 'X', 25-28, 'X', 30-31, 'X', 33-36 <S12>
A:Cross-references: UNIPARC:UPI0000177C39
R:Bernard, A.; Roels, H.; Lauwerys, R.; Witters, R.; Gielens, C.; Soumilion, A.; Van De
Clin. Chim. Acta 207, 239-249, 1992
A:Title: Human urinary protein 1: evidence for identity with the Clara cell protein and
A:Reference number: A56890; MWID:93009001; PMID:1395029
A:Accession: A56890
A:Molecule type: protein
A:Residues: 22-45 <BER>
A:Cross-references: UNIPARC:UPI0000036975
A:Experimental source: urine
A:Note: sequence extracted from NCBI backbone (NCBIP:119391)
R:Hay, J.G.; Daniel, C.; Chu, C.; Crystal, R.G.
Am. J. Physiol. 268, 565-575, 1995
A:Title: Human CC10 gene expression in airway epithelium and subchromosomal locus sugges
A:Reference number: I38397
A:Accession: I38397
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-91 <RES>
A:Cross-references: UNIPARC:UPI0000036901; EMBL:U01101; NID:g457932; PIDN:AAA8185.1; PI
C:Comment: This protein consists of two identical polypeptides linked by two disulfide b
C:Genetics:
A:Gene: CC10
C:Superfamily: uteroglobin
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-91/Product: Clara cell 10K protein #status experimental <MAT>

Query Match 51.6%; Score 16; DB 2; Length 91;
Best Local Similarity 22.2%; Pred. No. 1.1e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 16 CSSASATC 24

RESULT 46
UGMS
uteroglobin precursor - mouse
N:Alternate names: CC10; Clara cell 10K protein precursor; Clara cell secretory protein
C:Species: Mus musculus (house mouse)
C:Date: 03-May-1994 #sequence_revision 21-Jan-1997 #text_change 09-Jul-2004
C:Accession: A53025; A56656; I51925; S24783
R:Striip, B.R.; Huffman, J.A.; Bohinski, R.J.
Genomics 20, 27-35, 1994
A:Title: Structure and regulation of the murine Clara cell secretory protein gene.
A:Reference number: A53025; MWID:94292183; PMID:8020953
A:Accession: A53025
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-96 <STR>
A:Cross-references: UNIPROT:Q06318; UNIPARC:UPI00000011D0; GB:I24372; NID:g461147; PID:
R:Singh, G.; Katyal, S.L.; Brown, W.E.; Kennedy, A.L.
Exp. Lung Res. 19, 67-75, 1993
A:Title: Mouse Clara cell 10-KDa (CC10) protein: cDNA nucleotide sequence and molecular
A:Reference number: A56656; MWID:93178380; PMID:8440203
A:Accession: A56656
A:Molecule type: mRNA; protein
A:Residues: 1-96 <STN>
A:Cross-references: UNIPARC:UPI00000011D0; EMBL:X67702; NID:g49690; PIDN:CAA47936.1; PID
A:Experimental source: lung
A:Note: sequence extracted from NCBI backbone (NCBIP:126148)
A:Note: parts of this sequence, including the amino end of the mature protein, were conf
R:Margraf, L.R.; Finegold, M.J.; Stanley, L.A.; Major, A.; Naklins, H.K.; DeMayo, F.J.
Am. J. Respir. Cell Mol. Biol. 9, 231-238, 1993
A:Title: Cloning and tissue-specific expression of the cDNA for the mouse Clara cell 10K
A:Reference number: I51925; MWID:9400840; PMID:8398159
A:Accession: I51925
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-96 <RES>
A:Cross-references: UNIPARC:UPI00000011D0; GB:L04503; NID:g20213; PIDN:AAA03625.1; PID
A:Genetics:
A:introns: 19/1; 81/3
C:Complex: homodimer linked by two disulfide bonds
C:Superfamily: uteroglobin
C:Keywords: lung; steroid binding; uterus
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-96/Product: uteroglobin #status predicted <MAT>
F:24/Disulfide bonds: interchain (to 90) #status predicted
F:90/Disulfide bonds: interchain (to 24) #status predicted

Query Match 51.6%; Score 16; DB 1; Length 96;
Best Local Similarity 22.2%; Pred. No. 1.2e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 16 CSSASDTC 24

RESULT 47
A36581
polychlorinated biphenyl-binding protein precursor - rat
N:Alternate names: Clara cell 10K secretory protein
C:Species: Rattus norvegicus (Norway rat)
C:Date: 08-Mar-1991 #sequence_revision 08-Mar-1991 #text_change 09-Jul-2004
C:Accession: A36581; S10185; S21676
R:Nordlund-Moeller, L.; Andersson, O.; Ahlgren, R.; Schilling, J.; Gyllner, M.; Gustafar

J. Biol. Chem. 265, 12690-12693, 1990
A>Title: Cloning, structure, and expression of a rat binding protein for polychlorinated
A/Reference number: A36581; PMID:90324266; PMID:2115524
A/Accession: A36581
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-96 <NOR>
A/Cross-references: UNIPROT:P17559; UNIPARC:UPI0000112E8B; GB:J05536; NID:G206039; PIDN:
N1agen, G.; Wolf, M.; Karyal, S.L.; Singh, G.; Beato, M.; Suske, G.
Nucleic Acids Res. 18, 2939-2946, 1990
A>Title: Tissue-specific expression, hormonal regulation and 5'-flanking gene region of
A/Reference number: S10185; PMID:90272398; PMID:2349092
A/Accession: S10185
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-18 <HAG>
A/Cross-references: UNIPARC:UPI00001707CC; EMBL:X51318; NID:G55536; PIDN:CAA5701.1; PID
R/Umland, T.C.; Swaminathan, S.; Purey, W.; Singh, G.; Fletcher, J.; Sax, M.
J. Mol. Biol. 224, 441-448, 1992
A>Title: Refined structure of rat Clara cell 17 kDa protein at 3.0 Å resolution.
A/Reference number: S21676; PMID:92219263; PMID:1560460
A/Contents: annotation, X-ray crystallography, 3.0 angstroms
C/Superfamily: uteroglobin
F/24/Disulfide bonds: Interchain (to 90) #status experimental
F/30/Disulfide bonds: Interchain (to 24) #status experimental

Query Match 51.6%; Score 16; DB 2; Length 96;
Best Local Similarity 22.2%; Pred. No. 1.2e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 16 CSASSSDIC 24

RESULT 48
A10334
conserved hypothetical protein YPO2745 [imported] - *Yersinia pestis* (strain CO92)
C/Species: *Yersinia pestis*
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C/Accession: A10334
R/Parikh, J.; Wren, B.W.; Thomson, N.R.; Tibbail, R.W.; Holden, M.T.G.; Prentice, M.B.
demo-Ratnaga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.;
Jl. M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A>Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A/Reference number: AB0001; PMID:21470413; PMID:11586360
A/Accession: A10334
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-98 <KTR>
A/Cross-references: UNIPROT:Q8ZD47; UNIPARC:UPI00000CD967; GB:AL590842; PIDN:CAC92984.1;
C/Genetics:
A/Gene: YPO2745
C/Superfamily: Ca-dicarboxylate carrier protein

Query Match 51.6%; Score 16; DB 2; Length 98;
Best Local Similarity 22.2%; Pred. No. 1.2e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 8 CSAEETRAAC 16

RESULT 49
S76421
ferredoxin [2Fe-2S] - *Synechocystis* sp. (strain PCC 6803)
C/Species: *Synechocystis* sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C/Accession: S76421
R/Kaneko, T.; Sato, S.; Koriaki, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasud
DNA Res. 3, 109-136, 1996
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocysti*
S.
A/Reference number: S74322; PMID:97061201; PMID:8905231
A/Accession: S76421
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-103 <RAN>
A/Cross-references: UNIPROT:P74449; UNIPARC:UPI00000D7139; EMBL:D90915; GB:AB013339; NIT
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C/Superfamily: ferredoxin [2Fe-2S]; ferredoxin [2Fe-2S] homology
C/Keywords: 2Fe-2S; iron-sulfur protein; metalloprotein
F/28-84/Domains: ferredoxin [2Fe-2S] homology <FER>
F/43,48,51,83/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 51.6%; Score 16; DB 2; Length 103;
Best Local Similarity 22.2%; Pred. No. 1.2e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 83 CAAYATSDC 91

RESULT 50
J00863
hypothetical 11.6k protein - *Escherichia coli* retron Ec67
C/Species: *Escherichia coli* retron Ec67
C/Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C/Accession: J00863
R/Hsu, M.Y.; Inouye, M.; Inouye, S.
Proc. Natl. Acad. Sci. U.S.A. 87, 9454-9458, 1990
A>Title: Retron for the 67-base multicopy single-stranded DNA from *Escherichia coli*: a
A/Reference number: J00851; PMID:91067724; PMID:1701261
A/Accession: J00863
A/Molecule type: DNA
A/Residues: 1-104 <HSU>
A/Cross-references: UNIPROT:P21322; UNIPARC:UPI000013BF85; GB:M55249; NID:G145143; PIDN
A/Experimental source: E. coli strain Cl-1
C/Genetics:
A/Note: Insertion site is equivalent to 19 min of E. coli K12 genetic map
C/Superfamily: *Escherichia coli* retron Ec67 hypothetical 11.6k protein

Query Match 51.6%; Score 16; DB 2; Length 104;
Best Local Similarity 22.2%; Pred. No. 1.2e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 5 CSAEASHSC 13

Search completed: January 4, 2006, 16:10:08
Job time : 26.5217 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 ; Search time 97.9565 Seconds
(without alignments)
108.037 Million cell updates/sec

Title: US-09-932-322-3

Perfect score: 31
Sequence: 1 XXXXXXXXXXXXXXX 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 1000 summaries

Database : Uniprot_05.80:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	58.1	144	2	061KH6_DROME
2	18	58.1	155	2	P91214_CAEEL
3	18	58.1	269	2	07PR07_ANOGA
4	18	58.1	281	2	08GZ05_MAIZE
5	18	58.1	281	2	08LKK0_MAIZE
6	18	58.1	347	2	08Q850_ORYSA
7	18	58.1	416	2	09ST87_ORYSA
8	18	58.1	468	2	07XST8_ORYSA
9	18	58.1	470	2	09LFB4_ARATH
10	18	58.1	504	2	09XZ88_LEIMA
11	18	58.1	510	2	08S4F6_ARATH
12	18	58.1	513	2	07XZ21_XENLA
13	18	58.1	518	2	0640H3_XENLA
14	18	58.1	569	2	07OKT3_GIALA
15	18	58.1	684	2	07XV79_ORYSA
16	18	58.1	772	1	DLLA_BRABE
17	18	58.1	828	2	04WY29_ASPEU
18	18	58.1	1013	2	04SR96_TETNG
19	18	58.1	1031	2	04QFB5_LEIMA
20	18	58.1	1107	2	05CV70_CRYPV
21	18	58.1	1199	1	E7SBC_DROME
22	18	58.1	1355	1	E7SBA_DROME
23	18	58.1	1365	2	09VU79_DROME
24	18	58.1	1412	1	E7SBB_DROME
25	18	58.1	1793	2	09W451_DROME
26	18	58.1	1836	2	08MP07_DROME
27	18	58.1	1894	2	04OS15_LEIMA
28	18	58.1	2087	2	04S488_TETNG
29	17	54.8	50	2	09L981_VIBCH
30	17	54.8	71	2	04RA29_TETNG
31	17	54.8	96	2	08VD96_MESAU

32	17	54.8	108	2	04PAH6_USTWA	04PAH6_ustlago ma
33	17	54.8	108	2	09ZSP6_ARATH	09ZSP6_arabidopsis
34	17	54.8	115	2	07ZD14_DESVH	07ZD14_desulfovibr
35	17	54.8	117	2	06VZRI_CNPV	06VZRI_canarypox v
36	17	54.8	119	2	04TWL2_ASF	04TWL2_african swi
37	17	54.8	119	2	04TWL3_ASF	04TWL3_african swi
38	17	54.8	119	2	04TWL5_ASF	04TWL5_african swi
39	17	54.8	119	2	04TWL6_ASF	04TWL6_african swi
40	17	54.8	119	2	04TWL7_ASF	04TWL7_african swi
41	17	54.8	119	2	04TWL8_ASF	04TWL8_african swi
42	17	54.8	119	2	04TWL9_ASF	04TWL9_african swi
43	17	54.8	119	2	04TMO_ASF	04TMO_african swi
44	17	54.8	125	2	05YPN5_NOCPA	05YPN5_nocardi fa
45	17	54.8	133	2	04TZU4_TETNG	04TZU4_tetradon n
46	17	54.8	134	2	06ZRF5_HUMAN	06ZRF5_homo sapien
47	17	54.8	147	2	09CMG3_MOUSE	09CMG3_mus musculu
48	17	54.8	155	2	06R340_ORYSA	06R340_oryza sativ
49	17	54.8	155	2	04TWL4_ASF	04TWL4_african swi
50	17	54.8	157	1	C1025_HUMAN	081W50_homo sapien
51	17	54.8	157	1	C1025_MOUSE	09d772_mus musculu
52	17	54.8	157	2	04O6M1_LEIMA	04O6M1_leishmania
53	17	54.8	160	2	04J116_AZOVI	04J116_azobacter
54	17	54.8	163	2	09B1M0_GIALA	09B1M0_giardia lam
55	17	54.8	166	2	05TS94_HUMAN	05TS94_homo sapien
56	17	54.8	168	2	06K5J0_ORYSA	06K5J0_oryza sativ
57	17	54.8	170	2	08PRC3_XANAC	08PRC3_xanthomonas
58	17	54.8	171	2	04LIS2_GENTR	04LIS2_salmonella
59	17	54.8	171	2	093GN6_SALTU	093GN6_salmonella
60	17	54.8	172	2	05TS90_HUMAN	05TS90_homo sapien
61	17	54.8	172	2	06K2N4_ORYSA	06K2N4_oryza sativ
62	17	54.8	175	2	09VS07_DROME	09VS07_drosophila
63	17	54.8	183	2	05TS95_HUMAN	05TS95_homo sapien
64	17	54.8	186	1	C1025_BRABE	06C60_baccharomyc
65	17	54.8	191	2	05OZH7_BRABE	05OZH7_baccharomyc
66	17	54.8	199	2	07F588_CAEEL	07F588_caenorhabdi
67	17	54.8	202	2	06FX59_CANGA	06FX59_candida gla
68	17	54.8	204	2	06YU14_ORYSA	06YU14_oryza sativ
69	17	54.8	207	2	08N1Y5_HUMAN	08N1Y5_homo sapien
70	17	54.8	212	1	YN87_YEAST	PS3719_saccharomyc
71	17	54.8	214	2	053846_MYCTU	053846_mycobacteri
72	17	54.8	214	2	07U158_MYCTU	07U158_mycobacteri
73	17	54.8	222	2	04VS18_DROME	04VS18_drosophila
74	17	54.8	225	2	09LJ87_ARATH	09LJ87_arabidopsis
75	17	54.8	228	2	04VS54_DROME	04VS54_drosophila
76	17	54.8	244	2	09Y9Q2_AERPE	09Y9Q2_aeropyrum p
77	17	54.8	247	2	06BP27_DESHA	06BP27_debaryomyce
78	17	54.8	248	2	09USS9_SCHPO	09USS9_schizosacch
79	17	54.8	254	2	06WMS9_BRABE	06WMS9_brachyosto
80	17	54.8	257	2	06ZDL1_ORYSA	06ZDL1_oryza sativ
81	17	54.8	259	2	083212_TREPA	083212_treponema p
82	17	54.8	262	1	W0X3A_MAIZE	07U0V1_zea mays (m
83	17	54.8	265	1	W0X3B_MAIZE	06S313_zea mays (m
84	17	54.8	273	2	07SH84_ORYSA	07SH84_oryza sativ
85	17	54.8	281	2	057NZ1_SALCH	057NZ1_salmonella
86	17	54.8	281	2	05PHQ3_SALPA	05PHQ3_salmonella
87	17	54.8	281	2	08Z796_SALTU	08Z796_salmonella
88	17	54.8	281	2	08ZP77_SALTU	08ZP77_salmonella
89	17	54.8	283	2	07B973_MYCTU	07B973_mycobacteri
90	17	54.8	286	2	06GQD6_XENLA	06GQD6_xenopus lae
91	17	54.8	290	2	06A111_PROAC	06A111_propionibac
92	17	54.8	291	2	08W118_PROBI	08W118_vitis labru
93	17	54.8	293	1	XTM31_ARATH	P93046_arabidopsis
94	17	54.8	298	2	08Z599_ARATH	08Z599_arabidopsis
95	17	54.8	305	2	04LMB9_9BRBK	04LMB9_burholderi
96	17	54.8	305	2	067L98_SYTH	067L98_symblodact
97	17	54.8	306	2	07F8S0_ORYSA	07F8S0_oryza sativ
98	17	54.8	308	2	06ZJ81_BURMA	06ZJ81_burholderi
99	17	54.8	308	2	05GM13_CAEEL	05GM13_caenorhabdi
100	17	54.8	309	2	08GY34_ARATH	08GY34_arabidopsis
101	17	54.8	309	2	04NUR5_9DELT	04NUR5_aeromyxob
102	17	54.8	310	2	05N9V7_ORYSA	05N9V7_oryza sativ
103	17	54.8	311	2	06EP90_ORYSA	06EP90_oryza sativ
104	17	54.8	311	2	063SX3_BURPS	063SX3_burholderi

543	16	51.6	238	2	004364_ORYSA	004364_oryza sativ	616	16	51.6	284	2	Q96MV0_HUMAN	Q96MV0_homo sapien
544	16	51.6	238	2	067207_LYCES	067207_lycoopersico	617	16	51.6	285	2	Q7R3E3_GIALA	Q7R3E3_giardia lam
545	16	51.6	239	2	Q9NX17_HUMAN	Q9NX17_homo sapien	618	16	51.6	285	2	Q4P024_9DIPT	Q4P024_aeanaeromyxob
546	16	51.6	239	2	Q7R376_GIALA	Q7R376_giardia lam	619	16	51.6	285	2	Q9DGH7_CHICK	Q9DGH7_gallus gall
547	16	51.6	240	2	Q4P847_USTMA	Q4P847_ustiliago ma	620	16	51.6	286	2	Q7D0X6_AGRYS	Q7D0X6_agrobacteri
548	16	51.6	240	2	Q63MM3_BURPS	Q63MM3_burkholderi	621	16	51.6	286	2	Q9DBD2_MOUSE	Q9DBD2_mus musculu
549	16	51.6	242	2	Q51XD3_MAGGR	Q51XD3_magnaporthe	622	16	51.6	288	2	Q5Z6H0_ORYZA	Q5Z6H0_oryza sativ
550	16	51.6	243	2	Q8BJC3_MOUSE	Q8BJC3_mus musculu	623	16	51.6	289	2	Q9RXV6_DEIRA	Q9RXV6_deinococcus
551	16	51.6	243	2	Q93290_ORYLA	Q93290_oryza lat	624	16	51.6	291	2	Q7Q3X6_ANOXA	Q7Q3X6_anopheles g
552	16	51.6	244	2	Q5B708_EMENTI	Q5B708_aspergillus	625	16	51.6	291	2	Q4Z039_PLABE	Q4Z039_plasmodium
553	16	51.6	244	2	Q95ZL6_CAEBL	Q95ZL6_caenorhabdi	626	16	51.6	291	2	Q5VNP0_ORYSA	Q5VNP0_oryza sativ
554	16	51.6	246	2	Q62126_CAEBL	Q62126_caenorhabdi	627	16	51.6	295	2	Q07524_TROMA	Q07524_tropaeolum
555	16	51.6	246	2	Q4Z0Q4_PLABE	Q4Z0Q4_plasmodium	628	16	51.6	295	2	Q53NJ6_ORYSA	Q53NJ6_oryza sativ
556	16	51.6	246	2	Q6LND8_PHOPR	Q6LND8_photobacter	629	16	51.6	296	2	Q5OLC2_9DIPT	Q5OLC2_drosophila
557	16	51.6	246	2	Q8YDU4_BRUME	Q8YDU4_brucella me	630	16	51.6	296	2	Q5OLB1_DROBP	Q5OLB1_drosophila
558	16	51.6	247	2	Q966B6_CAEBL	Q966B6_caenorhabdi	631	16	51.6	296	2	Q5OLB0_DROBP	Q5OLB0_drosophila
559	16	51.6	247	2	Q4UT88_XANCP	Q4UT88_xanthomonas	632	16	51.6	296	2	Q5OLA9_9DIPT	Q5OLA9_drosophila
560	16	51.6	247	2	Q8PAD7_XANCP	Q8PAD7_xanthomonas	633	16	51.6	296	2	Q5OLA8_9DIPT	Q5OLA8_drosophila
561	16	51.6	247	2	Q923Z6_MOUSE	Q923Z6_mus musculu	634	16	51.6	296	2	Q5OLA7_9DIPT	Q5OLA7_drosophila
562	16	51.6	248	2	Q61HW9_CAEBR	Q61HW9_caenorhabdi	635	16	51.6	296	2	Q5OLA6_9DIPT	Q5OLA6_drosophila
563	16	51.6	248	2	Q4S4S0_CAEBL	Q4S4S0_caenorhabdi	636	16	51.6	296	2	Q5OLA5_9DIPT	Q5OLA5_drosophila
564	16	51.6	248	2	Q9D1R7_MOUSE	Q9D1R7_mus musculu	637	16	51.6	296	2	Q5OLA4_9DIPT	Q5OLA4_drosophila
565	16	51.6	249	2	Q414U5_GIBZE	Q414U5_gibberella	638	16	51.6	296	2	Q5OLA3_9DIPT	Q5OLA3_drosophila
566	16	51.6	249	2	Q5YPN1_NOCFA	Q5YPN1_nocardia fa	639	16	51.6	296	2	Q5OLA2_9DIPT	Q5OLA2_drosophila
567	16	51.6	249	2	Q916B6_XENIA	Q916B6_xenopus lae	640	16	51.6	297	1	TNR27_HUMAN	TNR27_homo sapien
568	16	51.6	252	2	Q95ZL7_CAEBL	Q95ZL7_caenorhabdi	641	16	51.6	297	2	Q5VYV0_HUMAN	Q5VYV0_homo sapien
569	16	51.6	252	2	Q9SZU5_ARYTH	Q9SZU5_arabidopsis	642	16	51.6	297	2	Q4VSR0_DROME	Q4VSR0_drosophila
570	16	51.6	252	2	Q8B9X9_SHEON	Q8B9X9_shewanella	643	16	51.6	298	2	Q5SHN3_CRYNE	Q5SHN3_cryptococcu
571	16	51.6	253	2	Q7YR96_BOVIN	Q7YR96_bos taurus	644	16	51.6	298	2	Q5H773_CRYNE	Q5H773_cryptococcu
572	16	51.6	253	2	Q9FYJ1_FESPR	Q9FYJ1_festuca pra	645	16	51.6	299	1	XTX32_ARYTH	XTX32_arythodopsis
573	16	51.6	254	2	Q5LUB1_SILPO	Q5LUB1_silicibacte	646	16	51.6	299	2	Q4V5X8_DROME	Q4V5X8_drosophila
574	16	51.6	254	2	Q5N1Y1_SYNP6	Q5N1Y1_synchococc	647	16	51.6	299	2	Q8GZFL_LYCES	Q8GZFL_lycoopersico
575	16	51.6	255	1	TNR9_HUMAN	Q07011_homo sapien	648	16	51.6	300	2	Q9HXY9_PSEBA	Q9HXY9_pseudomonas
576	16	51.6	256	2	Q4QP64_HABR18	Q4QP64_haemophilus	649	16	51.6	300	2	Q6Q960_LEIMA	Q6Q960_leishmania
577	16	51.6	258	2	Q6ON66_CAEBR	Q6ON66_caenorhabdi	650	16	51.6	301	2	Q6CXZ9_KLUDA	Q6CXZ9_kluyveromyc
578	16	51.6	258	2	Q06BG33_PARTE	Q06BG33_parmecium	651	16	51.6	301	2	Q6Z4N0_ORYSA	Q6Z4N0_oryza sativ
579	16	51.6	258	2	Q00946_PHYMN	Q00946_tetrahymena	652	16	51.6	302	2	Q5CND4_CRYHO	Q5CND4_cryptospori
580	16	51.6	258	2	Q82710_9ALPH	Q82710_gallid herp	653	16	51.6	302	2	Q7YXG0_CRYPV	Q7YXG0_cryptospori
581	16	51.6	259	2	Q51PM9_MAGGR	Q51PM9_magnaporthe	654	16	51.6	303	2	Q6W743_CRYPV	Q6W743_cryptospori
582	16	51.6	261	1	GSHI_MOUSE	P31315_mus musculu	655	16	51.6	303	2	Q6W744_CRYPV	Q6W744_cryptospori
583	16	51.6	261	1	Y118_HABIN	Q57097_haemophilus	656	16	51.6	306	2	Q4Q4F9_LEIMA	Q4Q4F9_leishmania
584	16	51.6	262	1	VME1_CVPSF	P09175_porcine tira	657	16	51.6	306	2	Q821D8_CHLACD	Q821D8_chlamydophi
585	16	51.6	264	1	GSHI_HUMAN	Q9H482_homo sapien	658	16	51.6	307	2	Q88XN9_PSEBK	Q88XN9_pseudomonas
586	16	51.6	264	1	Q67WK3_ORYSA	Q67WK3_oryza sativ	659	16	51.6	307	2	Q4WKG7_ASPERU	Q4WKG7_aspergillus
587	16	51.6	265	2	Q9ZEL5_CIOAB	Q9ZEL5_clostridium	660	16	51.6	307	2	Q9NZP0_9ICLI	Q9NZP0_parmecium
588	16	51.6	267	2	Q90257_PLDV	Q90257_fish lympho	661	16	51.6	308	2	Q86TW4_HUMAN	Q86TW4_homo sapien
589	16	51.6	268	1	HIS0_NEIGO	Q06758_neisseria g	662	16	51.6	308	2	Q7R414_GIALA	Q7R414_giardia lam
590	16	51.6	268	2	Q97AD5_THEVO	Q97AD5_thermoplasma	663	16	51.6	309	2	Q6CZH3_YARLI	Q6CZH3_yarrowia li
591	16	51.6	268	2	Q5F7L8_NEIG1	Q5F7L8_neisseria g	664	16	51.6	309	2	Q5U1Q0_ORYSA	Q5U1Q0_oryza sativ
592	16	51.6	268	2	Q9JTE4_NEIMA	Q9JTE4_neisseria m	665	16	51.6	310	1	Q91T37_LSDV	Q91T37_lumpy skin
593	16	51.6	268	2	Q9JYF0_NEIMB	Q9JYF0_neisseria m	666	16	51.6	310	1	TH14_HLSA	TH14_haloacteri
594	16	51.6	269	1	TH1D_STRCO	Q9ZB66_streptomyces	667	16	51.6	310	2	Q813Q5_PLAF7	Q813Q5_plasmodium
595	16	51.6	270	2	Q8WOTO_TTGCA	Q8WOTO_tigritopus c	668	16	51.6	310	2	Q9SZQ7_NOCAS	Q9SZQ7_nocardia as
596	16	51.6	272	2	Q4F8P7_FUGRU	Q4F8P7_fugu rubrip	669	16	51.6	310	2	Q7V249_PROMP	Q7V249_prochloroco
597	16	51.6	273	2	Q54QZ0_DICDI	Q54QZ0_dicyostelle	670	16	51.6	311	2	Q8BER3_XANAC	Q8BER3_xanthomonas
598	16	51.6	273	2	Q4TOT8_9SPHN	Q4TOT8_erythrobact	671	16	51.6	311	2	Q4WOC8_ASPERU	Q4WOC8_aspergillus
599	16	51.6	273	2	Q5RRU5_BBARE	Q5RRU5_brachydanio	672	16	51.6	312	2	Q5H6C2_XANBP	Q5H6C2_xanthomonas
600	16	51.6	275	2	Q44824_CAEBL	Q44824_caenorhabdi	673	16	51.6	312	2	Q72AR2_DBSYR	Q72AR2_desulfovibr
601	16	51.6	275	2	Q9CS10_MOUSE	Q9CS10_mus musculu	674	16	51.6	312	2	Q9DGH8_CHICK	Q9DGH8_gallus gall
602	16	51.6	276	1	FXL20_RAT	Q9GZH7_rattus norv	675	16	51.6	313	2	Q6M955_NEUCPO	Q6M955_neutropora
603	16	51.6	276	2	Q63Y45_BURPS	Q63Y45_arabidopsis	676	16	51.6	315	2	Q7ONT6_GIALA	Q7ONT6_giardia lam
604	16	51.6	277	2	Q6KY48_CHLRE	Q6KY48_burkholderi	677	16	51.6	315	2	Q7TTV7_SYNPX	Q7TTV7_synchococc
605	16	51.6	277	2	Q8KD78_MOUSE	Q8KD78_chlorobium	678	16	51.6	316	2	Q7NM80_GLOVI	Q7NM80_gloeobacter
606	16	51.6	277	2	Q06477_MOUSE	Q06477_mus musculu	679	16	51.6	317	2	Q53R75_HUMAN	Q53R75_homo sapien
607	16	51.6	278	1	YOX1_SCHPO	Q60098_schizosacch	680	16	51.6	318	2	Q5VYX9_HUMAN	Q5VYX9_homo sapien
608	16	51.6	278	2	Q8H8B1_ORYSA	Q8H8B1_oryza sativ	681	16	51.6	318	2	Q7NTH5_CHRVO	Q7NTH5_chromobacte
609	16	51.6	279	2	Q7SCA7_NEUCR	Q7SCA7_neutropora	682	16	51.6	319	2	Q5CUP4_CRYPV	Q5CUP4_cryptospori
610	16	51.6	279	2	Q511F7_MAGGR	Q511F7_magnaporthe	683	16	51.6	319	2	Q8WST3_CIOIN	Q8WST3_ciona intes
611	16	51.6	280	2	Q5B665_EMENTI	Q5B665_aspergillus	684	16	51.6	320	2	Q5CHH6_CRYHO	Q5CHH6_cryptospori
612	16	51.6	280	2	Q738T3_MYCPA	Q738T3_mycobacteri	685	16	51.6	320	2	Q7W116_BORBR	Q7W116_bordetella
613	16	51.6	281	2	Q6LSD9_ORYSA	Q6LSD9_oryza sativ	686	16	51.6	320	2	Q7VYZ0_BORPE	Q7VYZ0_bordetella
614	16	51.6	283	2	Q74Z18_ASHGO	Q74Z18_ashbya goss	687	16	51.6	321	2	Q59EP9_HUMAN	Q59EP9_homo sapien
615	16	51.6	283	2	Q9XKZ8_CERAE	Q9XKZ8_cercopithec	688	16	51.6	321	2	Q61EH2_CAEBR	Q61EH2_caenorhabdi

689	16	51.6	321	2	Q7W9Z1	BORPA	Q7w9z1	bordetella	762	16	51.6	379	2	Q043Z6	PICAN	Q043z6	piechia angu
690	16	51.6	323	2	Q5UIQ1	ORYSA	Q5uiq1	oryza sativ	763	16	51.6	381	2	Q59XB1	CANAL	Q59xb1	candida alb
691	16	51.6	323	2	Q9FRE1	ORYSA	Q9fre1	oryza sativ	764	16	51.6	382	1	FETUB	HUMAN	Q9uem5	homo sapien
692	16	51.6	323	2	Q9FRE3	ORYSA	Q9fre3	oryza sativ	765	16	51.6	382	2	Q59XF7	CANAL	Q59xf7	candida alb
693	16	51.6	325	2	Q8BJ16	MUSSE	Q8bj16	mus musculu	766	16	51.6	382	2	Q6DKS8	HUMAN	Q6dk58	homo sapien
694	16	51.6	327	1	TNR6	MOUSE	P25446	mus musculu	767	16	51.6	382	2	Q6GRB6	HUMAN	Q6grb6	homo sapien
695	16	51.6	327	2	Q8PFH9	BRMAJ	Q8pfh9	bradyrhizob	768	16	51.6	382	2	Q6ZAP3	ORYSA	Q6zaf3	oryza sativ
696	16	51.6	327	2	Q6GT31	MOUSE	Q6gt31	mus musculu	769	16	51.6	382	2	Q4RJ19	TERNG	Q4rj19	tetraodon n
697	16	51.6	329	2	Q7SBD9	NEOCR	Q7sbd9	neutrospora	770	16	51.6	383	1	BSC12	MOUSE	Q9z2es	mus musculu
698	16	51.6	331	2	Q8J359	TRSPA	Q8j359	treponema p	771	16	51.6	384	2	Q81Z08	PLAF7	Q81z08	plasmodium
699	16	51.6	332	2	Q4K6G9	PSEPS	Q4k6g9	pseudomonas	772	16	51.6	386	2	Q4SCA3	TERNG	Q4scA3	tetraodon n
700	16	51.6	333	1	PBX2	ARATH	Q73z60	arabidopsis	773	16	51.6	387	2	Q7WUB7	BORBP	Q7wub7	bordetella
701	16	51.6	333	2	Q73Z60	MYCPA	Q73z60	mycobacteri	774	16	51.6	387	2	Q7WPC9	BORBR	Q7wpc9	bordetella
702	16	51.6	334	2	Q84UD6	ARATH	Q84jd6	arabidopsis	775	16	51.6	388	2	Q64RH8	CHLS6	Q64rh8	chloratrachn
703	16	51.6	334	2	Q4ZNX1	PSBSY	Q4znx1	pseudomonas	776	16	51.6	388	2	Q631R1	BURPS	Q631r1	burkholderi
704	16	51.6	335	2	Q61S85	ORYSA	Q61sb5	oryza sativ	777	16	51.6	389	2	Q95018	CABEL	Q95018	caenorhabdi
705	16	51.6	336	2	Q5HEC0	XANOR	Q5hec0	xanthomonas	778	16	51.6	392	2	Q31277	RAT	Q31277	rattus norv
706	16	51.6	336	2	Q6LRP9	PHOPR	Q6lrf9	photobacter	779	16	51.6	394	2	Q4ICX6	GIBZE	Q4icx6	gibberella
707	16	51.6	336	2	Q5S808	SYNY3	Q5s808	synecocyst	780	16	51.6	394	2	Q54P67	DICDI	Q54p67	dictyostell
708	16	51.6	336	2	Q87WM6	PSBSM	Q87wm6	pseudomonas	781	16	51.6	394	2	Q88VH4	MOUSE	Q88vh4	mus musculu
709	16	51.6	337	2	Q5N412	SYNPE	Q5n412	synecococc	782	16	51.6	395	2	Q800K8	PAROL	Q800k8	paratrichity
710	16	51.6	338	2	Q4Q7J8	LEIMA	Q4q7j8	leishmania	783	16	51.6	396	2	Q4FYJ0	LEIMA	Q4fyj0	leishmania
711	16	51.6	338	2	Q67T05	SYNTH	Q67t05	syndiobacte	784	16	51.6	398	2	Q53JEN3	HUMAN	Q53jen3	homo sapien
712	16	51.6	339	2	Q7PZW5	ANOGA	Q7pzw5	anopheles g	785	16	51.6	398	2	Q567S1	HUMAN	Q567s1	homo sapien
713	16	51.6	339	2	Q6MNO7	BDBBA	Q6mng7	bdellovirib	786	16	51.6	400	1	BSC12	HUMAN	Q96g97	homo sapien
714	16	51.6	340	2	Q6LM46	PHOPR	Q6lm46	photobacter	787	16	51.6	400	2	Q61MR7	METMP	Q61mr7	mechanococ
715	16	51.6	341	1	PTAFR	MOUSE	Q62035	mus musculu	788	16	51.6	401	2	Q8BK84	MOUSE	Q8bk84	mus musculu
716	16	51.6	341	1	PTAFR	RAT	P46002	rattus norv	789	16	51.6	401	2	Q7G7C9	ORYSA	Q7g7c9	oryza sativ
717	16	51.6	341	2	Q4UTJ4	CORJK	Q4fj4	corynebacte	790	16	51.6	402	2	Q80WK4	MOUSE	Q80wk4	mus musculu
718	16	51.6	341	2	Q8UTX8	LSIDY	Q8jcx8	lumpy skin	791	16	51.6	402	2	Q5HOM2	XANOR	Q5hom2	xanthomonas
719	16	51.6	343	2	Q40148	LYCES	Q40148	lycopersico	792	16	51.6	404	2	Q86SB6	DROVI	Q86sb6	drosophi
720	16	51.6	344	2	Q7QZRS	GIALA	Q7qzrs	giardia lam	793	16	51.6	404	2	Q7W2C6	LEPIC	Q7w2c6	leptospira
721	16	51.6	344	2	Q8ZLT7	STRAM	Q8zlt7	streptomyce	794	16	51.6	404	2	Q8F9R4	LEPIN	Q8f9r4	leptospira
722	16	51.6	345	2	Q5J876	HUMAN	Q5j876	homo sapien	795	16	51.6	405	2	Q4V9L1	BRARE	Q4v9l1	bradyrhizob
723	16	51.6	345	2	Q73TJ0	MYCPA	Q73tj0	mycobacteri	796	16	51.6	408	2	Q6MNM7	BDBBA	Q6mnm7	bdellovirib
724	16	51.6	346	2	Q6ZAS3	ORYSA	Q6zas3	oryza sativ	797	16	51.6	408	2	Q8BFM1	MOUSE	Q8bfm1	mus muscu
725	16	51.6	347	2	Q8SYW5	DROME	Q8syw5	drosophi	798	16	51.6	408	2	Q8BMT1	MOUSE	Q8bmt1	mus musculu
726	16	51.6	347	2	Q5QD15	MOUSE	Q5qd15	mus musculu	799	16	51.6	409	2	Q8NK94	CABEL	Q8nk94	pholiora na
727	16	51.6	348	2	Q9RV18	DEIRA	Q9rv18	delnecocuss	800	16	51.6	411	2	Q4Y534	CABEL	Q4y534	caenorhabdi
728	16	51.6	348	2	Q8Z6B4	STRAM	Q8z6b4	streptomyce	801	16	51.6	413	2	Q5YUD1	NOCYA	Q5yud1	nocardia fa
729	16	51.6	350	2	Q7NGZ3	GLOVI	Q7ngz3	gloeobacter	802	16	51.6	414	2	Q7OH00	ANOGA	Q7oh00	anopheles g
730	16	51.6	351	2	Q8IGV7	DROME	Q8igv7	drosophi	803	16	51.6	416	1	TNR16	CHICK	TNR16	gallus gall
731	16	51.6	351	2	Q8MR04	DROME	Q8mr04	drosophi	804	16	51.6	416	2	Q6Z8K5	CABER	Q6z8k5	caenorhabdi
732	16	51.6	351	2	Q4RM41	TERNG	Q4rm41	tetraodon n	805	16	51.6	417	2	Q5VZP7	HUMAN	Q5vzp7	homo sapien
733	16	51.6	351	2	Q4TDL0	TERNG	Q4tld0	tetraodon n	806	16	51.6	418	2	Q9H599	HUMAN	Q9h599	homo sapien
734	16	51.6	355	1	RHOM	DROME	P20350	drosophi	807	16	51.6	418	2	Q21941	CABEL	Q21941	caenorhabdi
735	16	51.6	355	2	Q54OV7	DROME	Q54ov7	drosophi	808	16	51.6	418	2	Q9FPA0	ORYSA	Q9fpa0	oryza sativ
736	16	51.6	355	2	Q9XXJ5	CABEL	Q9xxj5	caenorhabdi	809	16	51.6	419	2	Q89SV5	BRMAJ	Q89sv5	bradyrhizob
737	16	51.6	355	2	Q94IQ4	ORYSA	Q94iq4	oryza sativ	810	16	51.6	419	2	Q4RO11	TERNG	Q4ro11	tetraodon n
738	16	51.6	356	2	Q8H8C7	ORYSA	Q8h8c7	oryza sativ	811	16	51.6	422	2	Q51AS0	BACFN	Q51as0	bacteroides
739	16	51.6	357	2	Q93SK6	MYXXA	Q93sk6	myxococcus	812	16	51.6	422	2	Q64Q68	BACFR	Q64q68	bacteroides
740	16	51.6	357	2	Q87685	SYVCZ	Q87685	chimpanzee	813	16	51.6	423	1	TNR19	HUMAN	TNR19	homo sapien
741	16	51.6	361	2	Q7OKX5	ANOGA	Q7okx5	anopheles g	814	16	51.6	423	2	Q5VZF8	HUMAN	Q5vzf8	homo sapien
742	16	51.6	362	2	Q5TW34	MACMU	Q5tm34	macaca mula	815	16	51.6	423	2	Q4LJU5	98BURK	Q4ljU5	98burk
743	16	51.6	362	2	Q5E0V5	VIBR1	Q5e0v5	vibrio flic	816	16	51.6	425	2	Q9VJ22	DROME	Q9vj22	drosophi
744	16	51.6	362	2	Q9X852	STRKO	Q9x852	streptomyce	817	16	51.6	425	2	Q7VZ43	PROMP	Q7vz43	prochloroco
745	16	51.6	363	2	Q7S9Y7	NEUCK	Q7s9y7	neutrospora	818	16	51.6	426	2	Q5DPT5	MOUSE	Q5dpt5	mus musculu
746	16	51.6	363	2	Q9VW94	DROME	Q9vw94	drosophi	819	16	51.6	427	2	Q7RXK3	NEUCK	Q7rxk3	neutrospora
747	16	51.6	363	2	Q9BJ44	LEIMA	Q9bj44	leishmania	820	16	51.6	428	2	Q5CRG0	CRYPV	Q5crg0	cryptospori
748	16	51.6	363	2	Q8H1T5	ABUTH	Q8h1t5	abutilon th	821	16	51.6	428	2	Q5CGW9	CRYHO	Q5cgw9	cryptospori
749	16	51.6	364	2	Q57W26	9TTRYP	Q57w26	tyrannosoma	822	16	51.6	429	2	Q60XZ1	CABER	Q60xz1	caenorhabdi
750	16	51.6	364	2	Q84XA2	MALXI	Q84xa2	malus xiao	823	16	51.6	432	2	Q4SLX3	TERNG	Q4slx3	tetraodon n
751	16	51.6	366	2	Q5NAP1	ORYSA	Q5nap1	oryza sativ	824	16	51.6	433	2	Q8WZ57	NEUCK	Q8wz57	neutrospora
752	16	51.6	366	2	Q4JW98	CORJK	Q4jw98	cozynebacte	825	16	51.6	433	2	Q7XLI3	ORYSA	Q7xli3	oryza sativ
753	16	51.6	366	2	Q4RJ83	TERNG	Q4rj83	tetraodon n	826	16	51.6	433	2	Q8LFT0	ARATH	Q8lft0	arabidopsis
754	16	51.6	369	2	Q6G4L0	BARBE	Q6g4l0	bartonella	827	16	51.6	433	2	Q9ZV84	ARATH	Q9zv84	arabidopsis
755	16	51.6	370	2	Q7OTU3	GIALA	Q7otus3	giardia lam	828	16	51.6	434	2	Q6BPN1	DEBHA	Q6bpn1	debaromyces
756	16	51.6	371	2	Q4LJTH	98BURK	Q4ljth	burkholderi	829	16	51.6	436	1	FXL20	HUMAN	FXL20	homo sapien
757	16	51.6	372	2	Q95TR6	DROME	Q95tr6	drosophi	830	16	51.6	436	1	Q8LUT8	HUMAN	Q8lUt8	homo sapien
758	16	51.6	377	1	BSC12	RAT	Q5fvl6	rattus norv	831	16	51.6	437	2	Q58T08	BRARE	Q58t08	bradyrhizob
759	16	51.6	377	2	Q9SUD8	HUMAN	Q9sud8	arabidopsis	832	16	51.6	437	2	Q4SUU7	TERNG	Q4suu7	tetraodon n
760	16	51.6	378	2	Q8QUV8	GVIRU	Q8quv8	infectiosis	833	16	51.6	438	2	Q9VSQ4	DROME	Q9vsq4	drosophi
761	16	51.6	378	2	Q8QUV8	GVIRU	Q8quv8	infectiosis	834	16	51.6	438	2	Q9VSQ4	DROME	Q9vsq4	drosophi

835	16	51.6	438	2	Q5BDG6_BOVIN	Q58d66 bos taurus	908	16	51.6	497	2	Q86BU3_BABO	Q86bu3 babeia bov
836	16	51.6	438	2	Q9DFV0_BRARE	Q9dfv0 bractyrdanio	909	16	51.6	499	2	Q4FKU6_9TRYP	Q4fkue trypanosoma
837	16	51.6	439	2	Q9DHF7_CLOAB	Q9dhf7 clostridium	910	16	51.6	500	2	Q20946_CAEBL	Q20946 caenorhabdi
838	16	51.6	441	2	Q4O7H1_LEIMA	Q4o7h1 leishmania	911	16	51.6	500	2	Q4ZOR8_PLABE	Q4z0r8 plasmodium
839	16	51.6	443	2	Q6EN36_ORYSA	Q6en36 oryza sativ	912	16	51.6	502	2	Q8CHV5_MOUSE	Q8chv5 mus musculu
840	16	51.6	445	2	Q818W6_GITALA	Q818w6 giardia lam	913	16	51.6	504	1	BTD_FUGRU	Q8av64 fugu rubrip
841	16	51.6	446	2	Q6FM66_CANGA	Q6fm66 candida gla	914	16	51.6	504	1	Q7SECI_NEUR	Q7seci1 neurospora
842	16	51.6	446	2	Q84S01_ORYSA	Q84s01 oryza sativ	915	16	51.6	504	2	Q6J1I8_9CAUD	Q6j1i8 trypanosoma
843	16	51.6	447	2	Q4ODU2_LEIMA	Q4odud2 leishmania	916	16	51.6	505	2	Q7RR38_PLAVO	Q7rr38 plasmodium
844	16	51.6	449	2	Q8OFV1_XENIA	Q8ofv1 xenopus lae	917	16	51.6	505	2	Q81756_9APIC	Q81756 plasmodium
845	16	51.6	451	2	Q4PBR2_USTMA	Q4pb2 ustilago ma	918	16	51.6	506	2	Q571F7_MOUSE	Q571f7 mus musculu
846	16	51.6	451	2	Q5TWM2_ANOCA	Q5tw2 anopheles g	919	16	51.6	510	2	Q968W3_9TRIP	Q968w3 trypanosoma
847	16	51.6	451	2	Q4RML6_TETNG	Q4rml6 tetraodon n	920	16	51.6	510	2	Q583L3_9TRYP	Q583l3 trypanosoma
848	16	51.6	453	2	Q4LUZ6_9BURK	Q4lu6 burkholderi	921	16	51.6	510	2	Q27072_TABSO	Q27072 taenia soli
849	16	51.6	453	2	Q6A0B7_MOUSE	Q6a0b7 mus musculu	922	16	51.6	512	1	VNN1_MOUSE	Q920t8 mus musculu
850	16	51.6	454	2	Q8ITU6_HUMAN	Q8itu6 homo sapien	923	16	51.6	513	1	INAR2_MOUSE	Q35664 mus musculu
851	16	51.6	454	2	Q7R3V9_GITALA	Q7r3v9 giardia lam	924	16	51.6	513	2	Q4PY68_LEIMA	Q4py68 leishmania
852	16	51.6	456	2	Q25242_LUCCU	Q25242 lucilia cup	925	16	51.6	513	2	Q92325_MOUSE	Q92325 mus musculu
853	16	51.6	457	2	Q7XPR7_ORYSA	Q7xpr7 oryza sativ	926	16	51.6	515	2	Q4PD21_USTMA	Q4pd21 ustilago ma
854	16	51.6	461	2	Q8A2G8_BACTN	Q8a2g8 bacteroides	927	16	51.6	515	2	Q70BY5_DROVI	Q70by5 dirosophila
855	16	51.6	461	2	Q5EGE1_BRARE	Q5ege1 brachydanio	928	16	51.6	516	2	Q57ZF4_9TRYP	Q57zf4 trypanosoma
856	16	51.6	462	2	Q7Z8U4_ASPOR	Q7z8u4 aspergillus	929	16	51.6	517	2	Q7VPZ8_CHLPP	Q7vpz8 chlamydia p
857	16	51.6	462	2	Q9C236_NEURC	Q9c236 neurospora	930	16	51.6	518	2	Q8N501_HUMAN	Q8n501 homo sapien
858	16	51.6	463	2	Q5LIU4_BACFN	Q5liu4 bacteroides	931	16	51.6	518	2	Q4UBB7_THEAN	Q4ubb7 theileria a
859	16	51.6	463	2	Q64ZY8_BACFR	Q64zy8 bacteroides	932	16	51.6	518	2	Q7DETO_MYCTU	Q7det0 mycobacteri
860	16	51.6	465	2	Q4RM22_TETNG	Q4rm22 tetraodon n	933	16	51.6	520	2	Q4FKK1_9TRYP	Q4fkk1 trypanosoma
861	16	51.6	466	2	Q757P9_ASHGO	Q757p9 ashbya goss	934	16	51.6	521	2	Q24867_HELPU	Q24867 helicobacte
862	16	51.6	466	2	Q7R7A2_PLAVO	Q7r7a2 plasmodium	935	16	51.6	523	2	Q9ZM39_HELPU	Q9zm39 helicobacte
863	16	51.6	466	2	Q5R9K3_FONPY	Q5r9k3 pongo pygma	936	16	51.6	523	2	Q7SYB1_BRARE	Q7syb1 brachydanio
864	16	51.6	467	2	Q82L86_STRAW	Q82l86 streptomyce	937	16	51.6	525	1	Y487_TREPA	Q83500 treponema p
865	16	51.6	468	1	TR10A_HUMAN	Q00220 homo sapien	938	16	51.6	526	2	Q7Q996_ANOCA	Q7q996 anopheles g
866	16	51.6	468	2	Q53Y72_HUMAN	Q53y72 homo sapien	939	16	51.6	526	2	Q6AQL6_DESPS	Q6aql6 desulfofale
867	16	51.6	468	2	Q9BB23_MACPA	Q9bbe3 macaca fasc	940	16	51.6	528	2	P71969_MYCTU	P71969 mycobacteri
868	16	51.6	468	2	Q4RSX7_MACPA	Q4rsx7 macaca fasc	941	16	51.6	528	2	Q7TV52_MYCBO	Q7tv52 mycobacteri
869	16	51.6	469	2	Q6GYV2_9INPA	Q6gyv2 influenza a	942	16	51.6	530	2	Q4IHU1_GIBZE	Q4ihu1 gibberella
870	16	51.6	469	2	Q6XUD6_IANAN	Q6xud6 influenza a	943	16	51.6	531	2	Q9VWZ8_DROME	Q9vwz8 drosophila
871	16	51.6	469	2	Q6XV53_9INPA	Q6xv53 influenza a	944	16	51.6	532	2	Q8P846_XANPC	Q8p846 xanthomonas
872	16	51.6	469	2	Q6XV61_9INPA	Q6xv61 influenza a	945	16	51.6	533	2	Q94G13_PHYPA	Q94g13 physcomitre
873	16	51.6	471	2	Q9VMG7_DROME	Q9vmg7 drosophila	946	16	51.6	533	2	Q72AF2_DESVH	Q72af2 desulfovibr
874	16	51.6	472	2	Q83MW3_TROWT	Q83mw3 tropheryma	947	16	51.6	535	2	Q5CYB9_PSRPV	Q5cyb9 cryptospori
875	16	51.6	472	2	Q83NLO_TROW8	Q83nlo tropheryma	948	16	51.6	536	2	Q4KXD6_PSRP5	Q4kxd6 pseudonias
876	16	51.6	476	1	VTDB_RABIT	P53789 oryctolagus	949	16	51.6	537	2	Q27005_TOXGO	Q27005 toxoplasma
877	16	51.6	476	2	Q7Q250_GITALA	Q7q250 giardia lam	950	16	51.6	538	2	Q5LEL4_CHLAB	Q5lel4 chlamydophi
878	16	51.6	478	2	Q70P55_9DELT	Q70p55 mellittangiu	951	16	51.6	539	2	Q4R3F5_MACFA	Q4r3f5 macaca fasc
879	16	51.6	479	2	Q61WA4_CAEBR	Q61wa4 caenorhabdi	952	16	51.6	540	1	Q4SUZ5_TETNG	Q4sus5 tetraodon n
880	16	51.6	479	2	Q82IU7_STRAW	Q82iu7 streptomyce	953	16	51.6	541	1	ATG322_CANGA	Q6fx22 candida gla
881	16	51.6	480	2	Q60YB4_CAEBR	Q60yb4 caenorhabdi	954	16	51.6	542	2	Q55NV1_CRYNE	Q55nv1 cryptococcu
882	16	51.6	480	2	Q25241_LUCCU	Q25241 lucilia cup	955	16	51.6	542	2	Q5KEL0_CRYNE	Q5kel0 cryptococcu
883	16	51.6	480	2	Q9XTG8_CAEBL	Q9xtg8 caenorhabdi	956	16	51.6	542	2	Q8N7C2_HUMAN	Q8n7c2 homo sapien
884	16	51.6	481	2	Q4RKF0_TETNG	Q4rkf0 tetraodon n	957	16	51.6	543	2	Q96192_CHLPP	Q96192 homo sapien
885	16	51.6	482	2	Q05932_PROHO	Q05932 prochloroth	958	16	51.6	543	2	Q9Z864_CHLPP	Q9z864 chlamydia p
886	16	51.6	482	2	Q5FW02_XENTR	Q5fw02 xenopus tro	959	16	51.6	544	2	Q5B195_DROME	Q5b195 pseudonias
887	16	51.6	483	2	Q80OK7_PAROL	Q80ok7 paratichthy	960	16	51.6	545	1	TREB_PSEAR	Q91165 pseudonias
888	16	51.6	484	2	Q4Z5Q4_PLABE	Q4z5q4 plasmodium	961	16	51.6	545	1	Q823Z7_CHLVC	Q823z7 chlamydophi
889	16	51.6	484	2	Q9W6F5_CHICK	Q9w6f5 gallus gall	962	16	51.6	548	2	Q4FKD2_9TRYP	Q4fk2 trypanosoma
890	16	51.6	486	2	Q64AN8_9ARCH	Q64an8 uncltured	963	16	51.6	550	2	Q8RLB6_BRADY	Q8rlb6 bradyrhizob
891	16	51.6	486	2	Q64WD4_CANGA	Q64wd4 candida gla	964	16	51.6	551	2	Q61GQ3_CABRR	Q61gq3 caenorhabdi
892	16	51.6	487	2	Q4RUT4_TETNG	Q4rut4 tetraodon n	965	16	51.6	551	2	Q7R3Y0_GITALA	Q7r3y0 giardia lam
893	16	51.6	487	2	Q8JUT8_XENIA	Q8jut8 xenopus lae	966	16	51.6	555	2	Q606Y8_MERCA	Q606y8 methyloccc
894	16	51.6	488	2	Q64A02_9ARCH	Q64a02 uncltured	967	16	51.6	557	2	Q88CUI_PSEBK	Q88cui pseudonias
895	16	51.6	489	2	Q64AD2_9ARCH	Q64ad2 uncltured	968	16	51.6	558	1	GPCL_HUMAN	Q83e11 pseudonias
896	16	51.6	489	2	Q6IDV7_DROPS	Q6idv7 drosophila	969	16	51.6	558	1	GPCL_RAT	P35053 rattus norv
897	16	51.6	490	2	Q4SSM9_TETNG	Q4ssm9 tetraodon n	970	16	51.6	558	2	Q53OM4_HUMAN	Q53om4 homo sapien
898	16	51.6	492	1	TISD_HUMAN	P47974 homo sapien	971	16	51.6	558	2	Q6P7O2_RAT	Q6p7g2 rattus norv
899	16	51.6	492	1	Q50NG3_ENTHI	Q50ng3 entamoeba h	972	16	51.6	559	2	Q8HZY6_MACPA	Q8hzy6 macaca fasc
900	16	51.6	493	1	LRC14_HUMAN	Q15048 homo sapien	973	16	51.6	561	2	Q5KEL1_CRYNE	Q5kel1 cryptococcu
901	16	51.6	493	2	Q57X17_9TRYP	Q57x17 trypanosoma	974	16	51.6	561	2	Q27007_TOXGO	Q27007 toxoplasma
902	16	51.6	493	2	Q569B5_RAT	Q569b5 rattus norv	975	16	51.6	566	2	Q580T6_9TRYP	Q580t6 trypanosoma
903	16	51.6	493	2	Q8VC16_MOUSE	Q8vc16 m leucine r	976	16	51.6	566	2	Q74BC4_GEOSL	Q74bc4 geobacter s
904	16	51.6	494	2	Q53TB4_HUMAN	Q53tb4 homo sapien	977	16	51.6	568	1	GAGJ_DROME	P21330 drosophila
905	16	51.6	494	2	Q80WK3_MOUSE	Q80wk3 mus musculu	978	16	51.6	571	2	Q68CN3_HUMAN	Q68cn3 homo sapien
906	16	51.6	495	2	Q6WZD6_9ARCH	Q6wzd6 uncltured	979	16	51.6	571	2	Q6H9L7_HUMAN	Q6h9l7 homo sapien
907	16	51.6	496	2	Q66671_GAWA	Q66671 equid herpe	980	16	51.6	571	2	Q8NBL0_HUMAN	Q8nbl0 homo sapien

```

981 16 51.6 571 2 Q6QH6_MOUSE Q6qhh6_mus_musculu
982 16 51.6 571 2 Q5EG07_IGTPU Q5eg07_ictalurus p
983 16 51.6 572 2 Q9J8B6_9NRC Q9j8b6_spodoptera
984 16 51.6 574 2 Q828T3_STRAM Q828t3_streptomyc
985 16 51.6 575 2 Q7QNV8_GIALA Q7qnv8_giardia lam
986 16 51.6 575 2 Q7XNU7_ORYSA Q7xnu7_oryza sativ
987 16 51.6 577 2 Q7UX93_RHODA Q7ux93_rhodospirill
988 16 51.6 578 2 Q7SG68_NEUCR Q7sg68_neutrospora
989 16 51.6 581 2 Q6WMM9_BOTCI Q6wmm9_botrytis ci
990 16 51.6 582 2 Q53PZ9_ORYSA Q53pz9_oryza sativ
991 16 51.6 587 1 SBU10_CAEEL Q93794_caenorhabdi
992 16 51.6 588 2 Q6GUE9_TOBAC Q6gues_nicotiana t
993 16 51.6 588 2 Q5UX66_BACRN Q5ux66_bacteroides
994 16 51.6 588 2 Q64XS9_BACRN Q64xs9_bacteroides
995 16 51.6 592 2 Q60018_PICAN Q60018_pichia angu
996 16 51.6 592 2 Q98H21_RHILQ Q98h21_rhizobium l
997 16 51.6 595 2 Q9C2K3_NEUCR Q9c2k3_neutrospora
998 16 51.6 595 2 Q5VXL7_HUMAN Q5vxl7_homo sapien
999 16 51.6 596 1 HMDH1_SOLTU P48020_solanum tub
1000 16 51.6 596 2 Q4K8Z7_PSEPS Q4k8z7_pseudomonas

```

ALIGNMENTS

RESULT 1
Q6IKH6_DROME PRELIMINARY; PRT; 144 AA.

```

AC 06IKH6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
HD C12427.
GN ORFNames=HDC12427;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14709175; DOI=10.1186/gb-2003-5-1-r3;
RA Hild M., Beckmann B., Haas S.A., Koch B., Solovjev V., Bueold C.,
RA Fellenberg K., Boutros M., Vingron M., Sauer F., Hottel J.D.,
RA Pao R.;
RT "An integrated gene annotation and transcriptional profiling approach
RT towards the full gene content of the Drosophila genome.";
RL Genome Biol. 5:RSEARCH0003.1-RESEARCH0003.17(2003).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ child party annotation (TPA) entry.
DR EMBL; BK002390; DAA03896.1; -; Genomic DNA.
SQ SEQUENCE 144 AA; 15277 MW; 1E4588BC183D5F2 CRC64;

```

Query Match 58.1%; Score 18; DB 2; Length 144;
Best Local Similarity 22.2%; Pred. No. 27;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 37 CATTASASC 45

RESULT 2
P91214_CAEEL PRELIMINARY; PRT; 155 AA.
AC P91214;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE C-type lectin protein 45.
GN Name=clec-45; ORFNames=F07C4.2;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;

```

OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Br1cel N2;
RC MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; U80023; AAG24036.1; -; Genomic DNA.
DR PIR; T28945; T28945.
DR Ensemble; F07C4.2; Caenorhabditis elegans.
DR Wormbase; WBGene001719; F07C4.2.
DR WormPep; F07C4.2; CE09201.
DR GO; GO:0005529; P:sugar binding; IEA.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; Lectin_C; 1.
DR SMART; SM00034; CLECT_1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
SQ Complete proteome; Lectin.
KW SEQUENCE 155 AA; 16285 MW; 956F0F4FDCB0F93 CRC64;

```

Query Match 58.1%; Score 18; DB 2; Length 155;
Best Local Similarity 22.2%; Pred. No. 28;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 92 CTTASSSSC 100

RESULT 3
Q7PRQ7_ANOGA PRELIMINARY; PRT; 469 AA.
AC Q7PRQ7;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000001657 (Fragment).
GN ORFNames=ENSANG00000001387;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxId=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST.
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

RC STRAIN=PEST.
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008647; BAA06779.3; -; Genomic DNA.
FT NON_TER 1
FT TER 269
SQ SEQUENCE 269 AA; 24228 MW; 25BBPF71FD71F1F2 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 269;
Best Local Similarity 22.2%; Pred. No. 34;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 228 CSSSTSSSC 236

RESULT 4

08GZ05_MAIZE PRELIMINARY; PRT; 281 AA.

ID 08GZ05_MAIZE PRELIMINARY; PRT; 281 AA.
AC 08GZ05;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE DREB-like protein.
GN Name=drebl;
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxId=4577;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Qin F., Li J., Zhao J., Chen S.-Y., Liu Q.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF448789; AAN7673.1; -; mRNA.
DR HSSP; 080337; ZGCC.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2; 1.
DR PRINTS; PR00367; ETHRSPELEMT.
DR ProDom; PD001423; TF_ERF; 1.
DR SMART; SM00380; AP2; 1.
SQ SEQUENCE 281 AA; 29346 MW; A0902A7FFE0AB978 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 281;
Best Local Similarity 22.2%; Pred. No. 35;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXC 12

Db 11 CSATTSSC 19

RESULT 5

08LKK0_MAIZE PRELIMINARY; PRT; 281 AA.

ID 08LKK0_MAIZE PRELIMINARY; PRT; 281 AA.
AC 08LKK0;
DT 01-OCT-2002 (TEMBLrel. 22, Created)
DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE DRE binding factor 2.
GN Name=dbf2;
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxId=4577;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=22058745; PubMed=12061899;
RA Kizis D., Pages M.;
RT "Maize DRE-binding proteins DBF1 and DBF2 are involved in rab17
RT regulation through the drought-responsive element in an ABA-dependent
RT pathway.";
RL Plant J. 30:679-689(2002).
DR EMBL; AF493799; AAM80485.1; -; mRNA.
DR HSSP; 080337; ZGCC.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2; 1.
DR PRINTS; PR00367; ETHRSPELEMT.
DR ProDom; PD001423; TF_ERF; 1.
DR SMART; SM00380; AP2; 1.

SQ SEQUENCE 281 AA; 29332 MW; CBE02A7FFE0AB979 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 281;
Best Local Similarity 22.2%; Pred. No. 35;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXC 12

Db 11 CSATTSSC 19

RESULT 6

080S50_9BETA PRELIMINARY; PRT; 347 AA.

ID 080S50_9BETA PRELIMINARY; PRT; 347 AA.
AC 080S50;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE UL38.
OS Pongine herpesvirus 4 (Chimpanzee cytomegalovirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxId=188763;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22421467; PubMed=12533697; DOI=10.1099/vir.0.18606-0;
RA Davison A.J., Dolan A., Akter P., Addison C., Dargan D.J.,
ALcendor D.J., McGeoch D.J., Hayward G.S.;
RT "The human cytomegalovirus genome revisited: comparison with the
RT chimpanzee cytomegalovirus genome.";
RL J. Gen. Virol. 84:17-28(2003).
DR EMBL; AF480884; AAM00688.1; -; Genomic DNA.
SQ SEQUENCE 347 AA; 39141 MW; 34B18107D501982 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 347;
Best Local Similarity 22.2%; Pred. No. 38;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXC 12

Db 312 CSSTSSSTC 320

09ST87_ORYSA PRELIMINARY; PRT; 416 AA.

ID 09ST87_ORYSA PRELIMINARY; PRT; 416 AA.
AC 09ST87;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE CAA303710.1 protein.
GN Name=q3037.10;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxId=4530;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=DNA;
RA Hong G., Chen Z.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ245900; CAB53483.1; -; Genomic DNA.
DR Gramene; 09ST87; -;
DR GO; GO:0004629; P:phospholipase C activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR000799; PI_PLX_X.
DR PROSITE; PS50007; PIPLC_X DOMAIN; 1.
SQ SEQUENCE 416 AA; 45736 MW; 332A261949FFE6DD CRC64;

Query Match 58.1%; Score 18; DB 2; Length 416;
Best Local Similarity 22.2%; Pred. No. 41;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 12
Db 34 CTASSASASC 42

RESULT 8

Q7XST8_ORYSA PRELIMINARY; PRT; 468 AA.

AC Q7XST8;

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)

DE OSUNBA0039K24.18 protein.

GN Name=OSUNBA0039K24.18;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

OX NCBI_TaxID=39947;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX PubMed=12447439; DOI=10.1038/nature01183;

RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,

RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,

RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,

RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Gan J.,

RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,

RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,

RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,

RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,

RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,

RA Han B.;

RT "Sequence and analysis of rice chromosome 4.";

RL Nature 420:316-320 (2002).

RE EMBL; AL606637; CAB01799.2; -; Genomic_DNA.

DR Gramene; Q7XST8; -;

DR GO; GO:0004629; P:phospholipase C activity; IEA.

DR GO; GO:0007242; P:intracellular signaling cascade; IEA.

DR InterPro; IPR000909; PI PLC_X.

DR PROSITE; PS50007; PIP2C_X DOMAIN; 1.

SO SEQUENCE 468 AA; 48671 MW; 49D661AA65338A3 CRC64;

Qy 4 CXXXXXXC 12

Db 34 CTASSASASC 42

RESULT 9

Q9LFB4_ARATH PRELIMINARY; PRT; 470 AA.

AC Q9LFB4;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DE Hypothetical protein F7J8_200.

GN Name=F7J8_200;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC euroside II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxID=3702;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Bevan M., Zimmermann W., Gruenisen A., Wambutt R., Kalicki J.,

RA Wohldmann P., Smith A., Bancroft I., Mewes H.W., Rudd S., Lemcke K.,

RA Mayer K.F.X.;

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA EU Arabidopsis sequencing project;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL137189; CAB69850.1; -; Genomic_DNA.

DR PIR; T45962; T45962.

DR GO; GO:0009058; P:biogenesis; IEA.

DR InterPro; IPR001296; Glyco_trans_1.

DR Pfam; PF00534; Glyco_transf_1; 1.

SO SEQUENCE 470 AA; 51797 MW; A7CB8DAA8EF5BC6 CRC64;

Qy 4 CXXXXXXC 12

Best local Similarity 58.1%; Score 18; DB 2; Length 470;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 23 CASSASASC 31

RESULT 10

Q9XZX8_LEIMA PRELIMINARY; PRT; 504 AA.

AC Q9XZX8;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DE Possible surface antigen.

GN Name=L302.01;

OS Leishmania major.

OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OX NCBI_TaxID=5664;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Friedlin;

RX MEDLINE=98146435; PubMed=9477341;

RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,

RA Smith D.F.;

RT "A physical map of the Leishmania major Friedlin genome.";

RL Genome Res. 8:135-145 (1998).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Friedlin;

RA Ivens A.C., Murphy L., Quail M., Harris D., Oliver K., Lawson D.,

RA Rajadream M.A., Barrell B.G.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL389894; CAC22669.1; -; Genomic_DNA.

DR InterPro; IPR006058; 2Fe2S_f4_BS.

DR InterPro; IPR006212; Furin_repeat.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR002049; Laminin_EGF.

DR SMART; SM00181; EGF; 3.

DR SMART; SM00261; FU; 5.

DR PROSITE; PS00197; 2FE2S_FERREDOXIN; UNKNOWN 3.

DR PROSITE; PS01248; LAMININ TYPE EGF; UNKNOWN 1.

SO SEQUENCE 504 AA; 52823 MW; B75450E3FE595263 CRC64;

Qy 4 CXXXXXXC 12

Db 161 CASTTASTC 169

RESULT 11

Q8S4F6_ARATH PRELIMINARY; PRT; 510 AA.

AC Q8S4F6;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)

DE Sulfolipid synthase (Hypothetical protein At5g01220).
 GN Name=SQD2; Synonyms=At5g01220;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxId=3702;
 (1)
 RN NUCLEOTIDE SEQUENCE.
 RP PubMed=11960029; DOI=10.1073/pnas.082696499;
 RA Yu B., Xu C., Benning C.;
 RT "Arabidopsis disrupted in SQD2 encoding sulfolipid synthase is
 impaired in phosphate-limited growth."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:5732-5737(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
 RA Onodera C.S., Quach H.L., Tang C.C., Toriumi M., Wong C., Wu H.C.,
 RA Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,
 RA Palm C.J., Shin P., Southwick A., Tripp M.G., Wu T., Davis R.W.,
 RA Becker J.R., Theologis A.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AA454354; AAM18913.1; -, mRNA.
 DR EMBL; BT005796; AA064198.1; -, mRNA.
 DR GO; GO:0009058; P:biosynthesis; IEA.
 DR InterPro; IPR001296; Glyco trans. 1.
 DR Pfam; PF00534; Glycos_transf_1; 1.
 DR Hypothetical protein.
 KW SEQUENCE 510 AA; 56630 MW; 20E57D318DBED68 CRC64;
 SQ

Query Match 58.1%; Score 18; DB 2; Length 510;
 Best Local Similarity 22.2%; Pred. No. 44;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXX 12
 DB 23 CXXXXXXX 31

RESULT 12
 Q7X21_XENLA PRELIMINARY; PRT; 513 AA.
 AC Q7X21;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE MGCS3520 protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OC NCBI_TaxId=8355;
 (1)
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Embryo;
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ushed T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Lounellano N.A., Peters G.J., Abramson R.D., Mullen S.J.,
 RA Bosa S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Mair M.A.,

RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 initiative."
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC046254; AAH46254.1; -, mRNA.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR003323; OTU.
 DR Pfam; PF02338; OTU; 1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS00802; OTU; 1.
 SQ SEQUENCE 513 AA; 55868 MW; A8734AD6E86FE8A CRC64;
 S

Query Match 58.1%; Score 18; DB 2; Length 513;
 Best Local Similarity 22.2%; Pred. No. 44;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXX 12
 DB 376 CXXXXXXX 384

RESULT 13
 Q640H3_XENLA PRELIMINARY; PRT; 518 AA.
 AC Q640H3;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE LOC494714 protein.
 GN Name=LOC494714;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OC NCBI_TaxId=8355;
 (1)
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Kidney;
 RC MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 initiative."
 RL Dev. Dyn. 225:384-391(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ushed T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Lounellano N.A., Peters G.J., Abramson R.D., Mullen S.J.,
 RA Bosa S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RN NUCLEOTIDE SEQUENCE.
 AC TISSUE-Kidney;
 RC Klein S., Gerhard D.S.;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC082654; AA82654.1; -; mRNA.
 DR InterPro; IPR003323; OTU.
 DR Pfam; PF02338; OTU; 1.
 DR PROSITE; PS50802; OTU; 1.
 SQ SEQUENCE 518 AA; 56209 MW; 2099AB07C6C93EC1 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 518;
 Best Local Similarity 22.2%; Pred. No. 45;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
 DB 381 CSATTAAC 389

RESULT 14
 Q7OX73_GITALA PRELIMINARY; PRT; 569 AA.
 AC Q7OX73;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE GLP 77 40692 38983.
 OS Giardia lamblia ATCC 50803.
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
 OC NCBI_TaxID=184922;
 (1)

NUCLEOTIDE SEQUENCE.
 RC STRAIN=WB C6;
 RX MEDLINE=21167845; PubMed=11104758; DOI=10.1074/jbc.M006589200;
 RA Abel E.S., Davids B.U., Robles L.D., Loflin C.E., Gillin F.D.,
 RA Chakrabarti R.;
 RT "Possible roles of protein kinase A in cell motility and excystation
 RT of the early diverging eukaryote Giardia lamblia.";
 RL J. Biol. Chem. 276:10320-10329(2001).
 RN [2]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=WB C6;
 RC STRAIN=WB C6;
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
 RA Olsen G.J., Sogin M.L.;
 RT "Draft sequence of the Giardia lamblia genome.";
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

CC EMBL; AACB0100056; EAA39878.1; -; Genomic DNA.
 DR GO; GO:0005489; F:electron transporter activity; IEA.
 DR GO; GO:0005506; F:iron ion binding; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001450; 4Fe4S_Fe_S_Bd.
 DR InterPro; IPR005127; 4Fe4S_VSP.
 DR Pfam; PF03302; VSP; 1.
 DR PRINTS; PRO0353; 4FE4SFPDOXIN.
 KM Electron transport; Transport.
 SQ SEQUENCE 569 AA; 59086 MW; EB0DF4E2B8CC3F20 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 569;
 Best Local Similarity 22.2%; Pred. No. 46;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
 DB 464 CSATTAAC 472

RESULT 15
 Q7XV79_ORYZA PRELIMINARY; PRT; 684 AA.
 ID Q7XV79;
 AC Q7XV79;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE OSUNB0076A22.6 protein.
 GN Name=OSUNB0076A22.6;
 GN Oryza sativa (Rice).
 OC Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=4530;
 (1)

NUCLEOTIDE SEQUENCE.
 RP MEDLINE=22337377; PubMed=12447439; DOI=10.1038/nature01183;
 RX Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
 RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Yang K., Yu S., Tang Y.,
 RA Weng Q., Zhang L., Lu Y., Mu Y., Lu Y., Zhang L.S., Yu Z., Fan D.,
 RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li Z., Hu H., Guan J.,
 RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
 RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
 RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
 RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
 RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
 RA Han B.;
 RT "Sequence and analysis of rice chromosome 4.";
 RL Nature 420:316-320(2002).
 RL EMBL; AL663016; CAD40795.3; -; Genomic DNA.
 DR Gramene; Q7XV79; -;
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0006740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN; 1.
 SQ SEQUENCE 684 AA; 75877 MW; 8CB1910F10409B14 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 684;
 Best Local Similarity 22.2%; Pred. No. 50;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
 DB 20 CSATTAAC 28

RESULT 16
 DDLA_BRARE STANDARD; PRT; 772 AA.
 ID DDLA_BRARE
 AC GED148; O57462;
 DT 13-SEP-2005 (Rel. 48, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Delta-like protein A precursor (Delta protein).
 GN Name=dlai;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxID:7955;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [mRNA], TISSUE SPECIFICITY, AND DEVELOPMENTAL
 RP STAGE.
 RX MEDLINE=98165392; PubMed=9425133;
 RA Appel B., Eisen J.S.;
 RT "Regulation of neuronal specification in the zebrafish spinal cord by
 RT Delta function.";
 RL Development 125:371-380 (1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Embryo;
 RG NIH - Zebrafish Gene Collection (ZGC) project;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP FUNCTION.
 RX PubMed=10572043;
 RA Riley B.B., Chiang M.-Y., Farmer L., Heck R.;
 RT "The delta gene of zebrafish mediates lateral inhibition of hair
 RT cells in the inner ear and is regulated by pax2.1.";
 RL Development 126:5669-5678 (1999).
 RN [4]
 RP INTERACTION WITH MIB, AND UBIQUITINATION.
 RX PubMed=15013799; DOI=10.1016/j.ydbio.2003.11.010;
 RA Chen W., Corliss P.C.;
 RT "Three modules of zebrafish Mind bomb work cooperatively to promote
 RT Delta ubiquitination and endocytosis.";
 RL Dev. Biol. 267:361-373 (2004).
 CC -1- FUNCTION: Acts as a ligand for Notch receptors and is involved in
 CC primary neurogenesis. Can activate Notch receptors, thereby
 CC playing a key role in lateral inhibition, a process that prevents
 CC the immediate neighbors of each nascent neural cell from
 CC simultaneously embarking on neural differentiation. In inner ear,
 CC it prevents adjacent cells from adopting the same cell fate.
 CC -1- SUBUNIT: Interacts with mib.
 CC -1- TISSUE SPECIFICITY: Type I membrane protein (By similarity).
 CC nervous system. It is expressed in overlapping regions with deltas
 CC (dlb) and delta2 (dl2), but differs the neural plate: it is
 CC expressed in patches of contiguous cells with dl2, while dlb is
 CC apparently confined to the scattered cells within those patches
 CC that differentiate as neurons. Expressed in hair cells of inner
 CC ear.
 CC -1- DEVELOPMENTAL STAGE: Initiated in the neuroectoderm before that of
 CC dld. In the developing trunk neural plate and neural tube, it is
 CC initiated in the epiblast prior to completion of gastrulation. At
 CC the 2- to 3-somite stage (10.5 hours) low levels are distributed
 CC throughout the trunk CNS, with cells expressing higher levels
 CC found in the medial and lateral regions of the neural plate. These
 CC regions correspond to the positions at which primary motoneurons
 CC and Rohon Bead neurons (Rbns) originate. Cells expressing high
 CC levels do not form contiguous domains. Rather, single cells or
 CC small clusters of several cells showing high expression are
 CC interspersed with cells having lower expression. Expressed is
 CC specific to the developing nervous system, and continues to be
 CC expressed broadly in the CNS throughout neurogenesis. Expressed in
 CC cells specified for neuronal fates. At 24 hours, and throughout
 CC later embryogenesis, it is broadly expressed in the spinal cord,
 CC suggesting that it is expressed by many types of cells. Expressed
 CC as neuronal specification occurs and is subsequently down-
 CC regulated in cells that have acquired specific neuronal fates.
 CC -1- PTM: Ubiquitinated by mib, leading to its endocytosis and
 CC subsequent degradation.
 CC -1- SIMILARITY: Contains 1 DSL domain.
 CC -1- SIMILARITY: Contains 8 EGF-like domains.
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 726.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC DR EMBL; AF030031; AAC41249.1; ALT SEQ; mRNA.
 CC DR EMBL; BC075742; AAH75742.1; -, mRNA.
 CC DR ZFIN; ZDB-GENE-980526-29; dlb.
 CC DR InterPro; IPR000152; Asx hydroxyl_S.
 CC DR InterPro; IPR001774; DSL.
 CC DR InterPro; IPR000742; EGF_2.
 CC DR InterPro; IPR001881; EGF_Ca_bd.
 CC DR InterPro; IPR001438; EGF_II.
 CC DR InterPro; IPR006209; EGF_1like.
 CC DR InterPro; IPR006210; IEGF.
 CC DR InterPro; IPR011651; MNFL.
 CC DR Pfam; PF01414; DSL; 1.
 CC DR Pfam; PF00008; EGF; 6.
 CC DR Pfam; PF07657; MNFL; 1.
 CC DR PRINTS; PR00010; EGFBLD.
 CC DR SMART; SM00051; DSL; 1.
 CC DR SMART; SM00181; EGF; 8.
 CC DR SMART; SM00179; EGF_CA; 6.
 CC DR PROSITE; PS00010; ASX_HYDROXYL; 2.
 CC DR PROSITE; PS1051; DSL; 1.
 CC DR PROSITE; PS00022; EGF_1; 8.
 CC DR PROSITE; PS01186; EGF_2; 8.
 CC DR PROSITE; PS50026; EGF_3; 7.
 CC DR PROSITE; PS01187; EGF_CA; 1.
 CC KM Calcium; Developmental protein; Differentiation; EGF-like domain;
 CC KW Glycoprotein; Neurogenesis; Notch signalling pathway; Repeat; Signal;
 CC KW Transmembrane; Ubl conjugation.
 CC FT SIGNAL 1 20 Potential.
 CC FT CHAIN 21 772 Delta-like protein A.
 CC FT TOPO_DOM 21 536 Extracellular (Potential).
 CC FT TRANSMEM 537 557 Potential.
 CC FT TOPO_DOM 558 -772 Cytoplasmic (Potential).
 CC FT DOMAIN 179 223 DSL.
 CC FT DOMAIN 225 257 EGF-like 1.
 CC FT DOMAIN 257 288 EGF-like 2.
 CC FT DOMAIN 290 328 EGF-like 3.
 CC FT DOMAIN 330 366 EGF-like 4; calcium-binding (Potential).
 CC FT DOMAIN 368 405 EGF-like 5.
 CC FT DOMAIN 407 443 EGF-like 6.
 CC FT DOMAIN 445 481 EGF-like 7.
 CC FT DOMAIN 483 519 EGF-like 8.
 CC FT COMPARS 697 746 Ser-rich.
 CC FT CARBOHYD 479 479 N-linked (GlcNAc...) (Potential).
 CC FT DISULFID 228 239 By similarity.
 CC FT DISULFID 232 245 By similarity.
 CC FT DISULFID 259 270 By similarity.
 CC FT DISULFID 265 276 By similarity.
 CC FT DISULFID 278 287 By similarity.
 CC FT DISULFID 294 306 By similarity.
 CC FT DISULFID 300 316 By similarity.
 CC FT DISULFID 318 327 By similarity.
 CC FT DISULFID 334 345 By similarity.
 CC FT DISULFID 339 354 By similarity.
 CC FT DISULFID 356 365 By similarity.
 CC FT DISULFID 372 383 By similarity.
 CC FT DISULFID 377 393 By similarity.
 CC FT DISULFID 395 404 By similarity.
 CC FT DISULFID 411 422 By similarity.
 CC FT DISULFID 416 431 By similarity.
 CC FT DISULFID 433 442 By similarity.
 CC FT DISULFID 449 460 By similarity.
 CC FT DISULFID 454 469 By similarity.
 CC FT DISULFID 471 480 By similarity.
 CC FT DISULFID 487 498 By similarity.
 CC FT DISULFID 492 507 By similarity.
 CC FT DISULFID 509 518 By similarity.
 CC SQ SEQUENCE 772 AA; 84969 MW; 716A014156938576 CRC64;
 CC -----
 CC Query Match 58.1%; Score 18; DB 1; Length 772;
 CC Best Local Similarity 22.2%; Pred. No. 52;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12

Db 740 CASASTSAC 748

RESULT 17

Q4WY29_ASPFU PRELIMINARY; PRT; 828 AA.

AC Q4WY29;

DT 13-SEP-2005 (TEMBLrel. 31, Created)

DT 13-SEP-2005 (TEMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TEMBLrel. 31, Last annotation update)

DE Repressor protein.

GN ORFNames=Afu3g14830;

OS Aspergillus fumigatus Af293.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

NCBI_TaxId=330879;

LN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Af293;

RA Niemman W., Pain A., Anderson M.J., Wortman J., Kim H., Stanley J.,

RA Artaya J., Berriman M., Abe K., Archer D.B., Bernedo C., Bennett J.,

RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,

RA Farman M., Fedorova N., Fedorova N., Fielding T.V., Fischer R.,

RA Foster N., Frazer A., Garcia J.L., Garcia M.J., Gobie A.,

RA Goldman G.H., Goni K., Griffith-Jones S., Gwilliam R., Haas B.,

RA Haas H., Harris D., Horluchi H., Huang J., Humphrey S., Jimenez J.,

RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,

RA Kumagai T., Latton A., Latge J.-P., Li W., Lord A., Lu C.,

RA Majeros W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,

RA Mouyia I., Mulligan S., Murphy L., O'Neill S., Paulsen I.,

RA Penava M.A., Perera M., Price C., Pritchard B.L., Quail M.A.,

RA Rabinovitch E., Rawlins N., Rajandream M.-A., Reichard U.,

RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,

RA Roming C.M., Rutter S., Salzberg S.L., Sanchez M., Squares S.,

RA Sanchez-Ferreiro J.C., Saunders D., Seeger K., Squares R., Squares S.,

RA Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,

RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,

RA Machida M., Hall N., Barrell B., Denning D.W.,

RT "Genomic sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus."

RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.

CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL; AF01000002; EAL92104.1; -!- Genomic DNA.

CC DR EMBL; AF01000002; 2B1B831C5E4FF0C6 CRC64;

SQ SEQUENCE 828 AA; 92421 MW; 2B1B831C5E4FF0C6 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 828;

Best Local Similarity 22.2%; Pred. No. 54;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12

Db 57 CTTTTTTC 65

RESULT 18

Q4SR96_TETNG PRELIMINARY; PRT; 1013 AA.

AC Q4SR96;

DT 13-SEP-2005 (TEMBLrel. 31, Created)

DT 13-SEP-2005 (TEMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TEMBLrel. 31, Last annotation update)

DE Chromosome 11 SCAR14528, whole genome shotgun sequence.

DE (Fragment).

GN ORFNames=GSTENG00014015001;

OS Tetraodon nigroviridis (Green puffer).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;

OC Tetraodontidae; Tetraodontidae; Tetraodon.

NCBI_TaxId=99883;

LN [1]

RP NUCLEOTIDE SEQUENCE.

RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,

RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,

RA Niclaud S., Jaffe D., Fisher S., Lutfalla G., Dosset C., Segreus B.,

RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,

RA Anthouard V., Jubin C., Castellic V., Katinka M., Vacherie B.,

RA Biemont C., Skalli Z., Catolico L., Poulin J., De Bernardis V.,

RA Craud C., Duprat S., Brotier P., Coutanceau J.P., Gouzy J.,

RA Parra G., Lardier G., Chappie C., McKernan K.J., McEwan P., Bosak S.,

RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,

RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,

RA Landet V., Schachter V., Quetier F., Sautin W., Scarpelli C.,

RA Winkler P., Lander E.S., Weissbach J., Roest Croillins H.,

RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype."

RL Nature 431:946-957(2004).

LN [2]

RP NUCLEOTIDE SEQUENCE.

RG Genoscope; Whitehead Institute Centre for Genome Research;

RL Submitted (Feb-2004) to the EMBL/GenBank/DBJ databases.

CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL; CAAB0101528; CAP96836.1; -!- Genomic DNA.

DR InterPro; IPR000175; Na/ntiran_symport.

DR Pfam; PF00209; SNF; 2.

DR PRINTS; PR00176; NANUSMPORT.

DR ProDom; PD000448; Na/ntiran_symport; 2.

DR PROSITE; PS00610; NA_NEUROTRAN_SYM_1; 1.

DR PROSITE; PS00754; NA_NEUROTRAN_SYM_2; UNKNOWN_1.

DR PROSITE; PS0267; NA_NEUROTRAN_SYM_3; 2.

KT Symport; Transmembrane; Transport.

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

SQ SEQUENCE 1013 AA; 110756 MW; 3D35CF87787652D CRC64;

Query Match 58.1%; Score 18; DB 2; Length 1013;

Best Local Similarity 22.2%; Pred. No. 58;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12

Db 728 CSASTSAC 726

RESULT 19

Q4QFB5_LEIMA PRELIMINARY; PRT; 1031 AA.

AC Q4QFB5;

DT 13-SEP-2005 (TEMBLrel. 31, Created)

DT 13-SEP-2005 (TEMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TEMBLrel. 31, Last annotation update)

DE Hypothetical protein.

GN ORFNames=LmjF15.0700;

OS Leishmania major.

OC Eukaryota; Euzoaria; Kinetoplastida; Trypanosomatidae; Leishmania.

NCBI_TaxId=5664;

LN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Friedlin;

RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,

RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,

RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.,

RL Submitted (Jun-2005) to the EMBL/GenBank/DBJ databases.

DR EMBL; CT005254; CAJ03294.1; -!- Genomic DNA.

DE Hypothetical protein.

KT Hypothetical protein.

SQ SEQUENCE 1031 AA; 112673 MW; 65DB334CA5ECBA CRC64;

Query Match 58.1%; Score 18; DB 2; Length 1031;

Best Local Similarity 22.2%; Pred. No. 59;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXC 12
DB 791 CXXXXSSSSAC 799

RESULT 20

Q5CV70_CRYPV PRELIMINARY; PRT; 1107 AA.
AC Q5CV70;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, last sequence update)
DE Hypothetical protein.
GN ORFNames=cg48_5210;
OS Crypsosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Crypsosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Iowa type II;
RX PubMed=15044751; DOI=10.1126/science.1094786;
RA Abrahamson M.S., Templeton T.J., Enomoto S., Abraham J.E., Zhu G.,
Rancic C.A., Deng M., Liu C., Widmer G., Tzipori S., Buck G.A., Xu P.,
R Bankler A.T., Dear P.H., Konfortov B.A., Spriggs H.F., Iyer L.,
R Anantharaman V., Aravind L., Kapur V.;
RT "Complete genome sequence of the apicomplexan, Cryptosporidium
parvum.";
RL Science 304:441-445 (2004).
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL: AABE0100003; EAK89643.1; -; Genomic_DNA.
DR InterPro: IPR012283; 6PGD_C core.
DR InterPro: IPR011989; ARM-like.
DR InterPro: IPR012293; Globin-related.
DR InterPro: IPR001524; Glyco_Hydro_6.
DR InterPro: IPR012287; Homodomain-rel.
DR InterPro: IPR001093; IMPH/GMPase.
DR InterPro: IPR008976; PLAT IM2.
DR InterPro: IPR000783; RNAPol_RPB5.
DR InterPro: IPR009033; Ypu_Cyt.
DR InterPro: IPR011991; Wing_hix_DNA_bd.
DR Pfam: PF00478; IMPDH_1.
KW Hypothetical protein.
SQ SEQUENCE 1107 AA; 127343 MW; 4BAD9AEB36ECB2AF CRC64;

Query Match 58.1%; Score 18; DB 2; Length 1107;
Best Local Similarity 22.2%; Pred. No. 60;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXC 12
DB 185 CXXXXSSSSC 193

RESULT 21

E75BC DROME
AC P17671; O810S1; O810S3; PRT; 1199 AA.
DT 01-AUG-1990 (Rel. 15, Created)
DT 13-SEP-2005 (Rel. 48, last sequence update)
DT 13-SEP-2005 (Rel. 48, last annotation update)
DE Ecdysone-induced protein 75B isoforms C/D (E75-A).
GN Name=E75B; Synonyms=NR1D3; ORFNames=CG8127;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

RN [1]
RC NUCLEOTIDE SEQUENCE [MRNA], AND ALTERNATIVE SPLICING.
RP STRAIN=Cancon-S;
RX MEDLINE=90249727; PubMed=2110921;
RA Segaves W.A., Hognes D.S.;
RT "The E75 ecdysone-inducible gene responsible for the 75B early puff in
RT Drosophila encodes two new members of the steroid receptor
RT superfamily.";
RL Genes Dev. 4:204-219 (1990).
RN [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abmayri A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck Y., Brokstein P., Brotier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.B., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenton J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lei Z.,
RA Lascko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nishikura D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Stryker R., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Sylvestre R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
RN [3]
RP GENOME REANNOTATION, AND ALTERNATIVE SPLICING.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.B.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Betencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).
RN [4]
RP FUNCTION, DEVELOPMENTAL STAGE, AND INDUCTION.
RX MEDLINE=94038699; PubMed=8223281;
RA Huot F., Ruiz C., Richards G.;
RT "Puffs and PCR: the in vivo dynamics of early gene expression during
RT ecdysone responses in Drosophila.";
RL Development 118:613-627 (1993).
CC -1- FUNCTION: Implicated in the regulation of ecdysone-triggered gene
hierarchies. Probably plays a key role in mediating the regulation

CC of the larval molt by 20-OH-ecdysone.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=C; Synonyms=E75A;
 CC IsoId=P1671-1; Sequence=Displayed;
 CC Name=A; Synonyms=E75B;
 CC IsoId=P1672-1; Sequence=External;
 CC Name=B; Synonyms=E75C;
 CC IsoId=P13055-1; Sequence=External;
 CC Name=D;
 CC IsoId=P1671-2; Sequence=VSP 014915; VSP 014916;
 CC Note=No experimental confirmation available;
 CC -1- DEVELOPMENTAL STAGE: In mid instar larvae salivary glands, low
 CC basal levels are observed in puff stage 1. Levels increase in late
 CC larvae from puff stages 3-10, then decrease abruptly at stage 11.
 CC In prepupae, isoform C is the predominant form during the
 CC transition between puff stages 18-19. At puff stage 1, expression
 CC is also present in the gut. By stage 3 it is present in the wing
 CC disks, Malpighian tubules and the fat body. At stage 11,
 CC expression is only present in the gut and wing disks.
 CC -1- INDUCTION: The expression of this protein is developmentally
 CC regulated and is correlated with the 20-OH-ecdysone induced
 CC activity of puff 75B.
 CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NR1
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 nuclear receptor DNA-binding domain.
 CC -1- CAUTION: Ref. 1 sequence differs from that shown due to a
 CC frameshift in position 1164.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL, X51548; CAA35923.1; ALT FRAME: mRNA.
 CC EMBL, AE003522; AAN11687.1; -; Genomic DNA.
 CC EMBL, AE003522; AAN11689.1; -; Genomic DNA.
 CC PIR, A34598; A34598.
 CC HSSP; P20393; 1GA5.
 CC SMR; P1671; 240-323.
 CC TRANSFAC; T01367; -;
 CC PLAYBASE; FBgn000568; E1p75B.
 CC GO; GO:0004879; P.ligand-dependent nuclear receptor activity; NAS.
 CC GO; GO:0018990; P.ecdysis (senu insects); IMP.
 CC GO; GO:0035072; P.ecdysone-mediated induction of salivary gla. .; NAS.
 CC GO; GO:0007553; P.regulation of ecdysteroid metabolism; IMP.
 CC InterPro; IPR001628; Hrmn_rcpt_DNA_bd.
 CC InterPro; IPR000536; Hrmn_rcpt_lig_bd.
 CC InterPro; IPR001723; Stdhnm_receptor.
 CC InterPro; IPR000324; Vldt_receptor.
 CC Pfam; PF00104; Hormone_recep; 1.
 CC Pfam; PF00105; zf-C4; 1.
 CC PRINTS; PR000398; STRDHORMONER.
 CC PRINTS; PR00047; STROIDRINGER.
 CC PRINTS; PR00350; VITAMINDR.
 CC ProDom; PD000035; Znf_C4steroid; 1.
 CC SMART; SM00430; HOL1; 1.
 CC SMART; SM00399; znf_C4; 1.
 CC PROSITE; PS00031; NUCLEAR_REC_DBD_2; 1.
 CC PROSITE; PS01030; NUCLEAR_REC_DBD_2; 1.
 CC Alternative splicing; Developmental protein; DNA-binding;
 CC Metal-binding; Nuclear protein; Receptor; Transcription;
 CC Transcription regulation; Zinc; Zinc-finger.
 CC TRANSCRIPTION regulation; Zinc; Zinc-finger.
 CC FT DNA_BIND 242 318
 CC FT ZN_RING 245 265 NR C4-type.
 CC FT ZN_RING 282 306 NR C4-type.
 CC FT REGION 380 590 Ligand-binding (Potential).
 CC FT COMPIAS 139 195 Gln/His-rich.
 CC FT COMPIAS 721 756 Gln-rich.
 CC FT COMPIAS 900 1103 Ser-rich.
 CC FT VARSPLIC 1 293 Missing (in isoform D).

FT FTID=VSP 014915.
 FT VARSPLIC 294 315 NNNRCYCKKCKIAVGMGRDA -> MGEELPIKIGILKGN
 FT VARSPLIC (in isoform D).
 FT /FTID=VSP 014916.
 SQ SEQUENCE 1199 AA; 128534 MW; E29372F5E295F595 CRC64;
 Query Match 58.1%; Score 18; DB 1; Length 1199;
 Best Local Similarity 22.2%; Pred. No. 63;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 4 CXXXXXXX 12
 Db 1070 CXXXXXXX 1078
 RESULT 22
 ID E75BA DROME STANDARD; PRT; 1355 AA.
 AC E75BA DROME 081052;
 DT 01-AUG-1990 (Rel. 15 Created)
 DT 13-SEP-2005 (Rel. 48, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Ecdysone-induced protein 75B isoform A (E75-B).
 GN Name=E1p75B; Synonyms=NR1D3; ORFNames=C8127;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 [1]
 RN RP NUCLEOTIDE SEQUENCE [MRNA], AND ALTERNATIVE SPLICING.
 RC STRAIN=canon-S;
 RX MEDLINE=90249727; PubMed=2110921;
 RT "The E75 ecdysone-inducible gene responsible for the 75B early puff in
 RT Drosophila encodes two new members of the steroid receptor
 RT superfamily.";
 RL Gene Dev. 4:204-219(1990).
 [2]
 RN RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Ceinker S.E., Holt R.A., Iwano C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Chang Q., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt D., Nelson C.R., Milos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davaport L.B., Davies A.,
 RA de Pablos B., Delcher A., Deng Z., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.B., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasok P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy J., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).
 [3]
 RN GENOME REANNOTATION, AND ALTERNATIVE SPLICING.
 RP MEDLINE:22426069; PubMed:12537572.
 RA Mista S., Crosby M.A., Mungall C.U., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Gelinker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.,
 RA "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review,"
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 [4]
 RN FUNCTION, DEVELOPMENTAL STAGE, AND INDUCTION.
 RP MEDLINE:94038699; PubMed:8223281;
 RA Hueb F., Ruiz C., Richards G.,
 RT "Puffs and PCR: the in vivo dynamics of early gene expression during
 RT ecdysone responses in *Drosophila*,"
 RL Development 118:613-627(1993).
 CC -1- FUNCTION: Implicated in the regulation of ecdysone-triggered gene
 CC hierarchy. Probably plays a key role in mediating the regulation
 CC of the larval molt by 20-OH-ecdysone.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=4;
 CC Name=A; Synonyms=E75B;
 CC IsoId=PI1672-1; Sequence=Displayed;
 CC Name=C; Synonyms=E75A;
 CC IsoId=PI1671-1; Sequence=External;
 CC Name=B; Synonyms=E75C;
 CC IsoId=PI1055-2; Sequence=External;
 CC Name=D;
 CC IsoId=PI1671-2; Sequence=External;
 CC Note=No experimental confirmation available;
 CC -1- DEVELOPMENTAL STAGE: In mid instar larvae salivary glands, levels
 CC are low during puff stage 1, increase during puff stages 2-4 and
 CC diminish from stage 5 onwards. In prepupae, isoform A is the
 CC predominant form during puff stage 19 and the transition to stage
 CC 20. By stage 3 it is present in the gut, Malpighian tubules and
 CC the fat body, levels persist beyond stage 11.
 CC -1- INDUCTION: The expression of this protein is developmentally
 CC regulated and is correlated with the 20-OH-ecdysone induced
 CC activity of puff 75B.
 CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NRI
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 nuclear receptor DNA-binding domain.
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 1320.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL: X51549; CAA35924.1; ALT_FRAME; mRNA.
 CC EMBL: AE003522; AAN11688.1; -; Genomic_DNA.
 CC PIR: B34598; B34598.
 CC HSSP: P20393; 1A6Y.
 CC SMR: P1672; 415-480.
 CC TRASPAC: T01368; -;
 CC Flybase: FBgn0000568; E1P75B.
 CC GO: GO:0004879; F:11gand-dependent nuclear receptor activity; NAS.
 CC GO: GO:0019990; P:ecdysis (sensu insecta); IMP.
 CC GO: GO:0035072; P:ecdysone-mediated induction of salivary gla. .; NAS.

DR GO: GO:0007553; P:regulation of ecdysteroid metabolism; IMP.
 DR InterPro: IPR001628; Hrm rcpt DNA bd.
 DR InterPro: IPR000536; Hrm rcpt_lig bd.
 DR InterPro: IPR000354; Involucrin_rpt.
 DR InterPro: IPR003079; KOR_receptor.
 DR InterPro: IPR001723; Stdhmg_receptor.
 DR Pfam: PF00104; Hormone_recep; 1.
 DR Pfam: PF00904; Involucrin; 1.
 DR Pfam: PF00105; zf-C4; 1.
 DR PRINTS: PR01293; KORNUCCEPTR.
 DR PRINTS: PR00396; STRDHOMNER.
 DR PRINTS: PR00047; STROIDFINGER.
 DR ProDom: PD000035; ZnfC4steroid; 1.
 DR SMART: SM00430; HOL1; 1.
 DR SMART: SM00399; ZnfC4; 1.
 DR PROSITE: PS51030; NUCLEAR_REC_DBD_2; 1.
 DR Alternative splicing; Developmental protein; DNA-binding;
 KW Metal-binding; Nuclear protein; Receptor; Transcription;
 KW Transcription regulation; Zinc; Zinc-finger.
 FT DNA_BIND 384 474 Nuclear receptor.
 FT ZN_FING 387 421 NR C4-type 1; degenerate.
 FT ZN_FING 438 457 NR C4-type 2.
 FT COMPBIAS 6 324 Gin-rich.
 FT COMPBIAS 877 912 Gin-rich.
 FT COMPBIAS 1056 1259 Ser-rich.
 FT CONFLICT 206 206 S -> C (in Ref. 1).
 FT CONFLICT 244 244 L -> L (in Ref. 1).
 FT CONFLICT 244 244 L -> L (in Ref. 1).
 SQ SEQUENCE 1355 AA; 147171 MW; 6EF19BACEC56225 CRC64;
 Query Match 58.1%; Score 18; DB 1; Length 1355;
 Best Local Similarity 22.2%; Pred. No. 65;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 4 CXXXXXXKC 12
 DB 1226 CXXXXXXKC 1234
 RESULT 23
 ID Q9VU79 DROME PRELIMINARY; PRT; 1365 AA.
 AC Q9VU79;
 DT 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
 DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)
 DE CG10231-PA.
 GN Name=Pdel1; ORFNames=CG10231;
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 OC NCBI_TaxId=7227;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE:20196006; PubMed:10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Gelinker S.E., Holt R.A., Evans C.A., Gockyne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blake J.R., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.W., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Foster K., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abril J.F., Abdayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Bence P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Bortova K.C., Busan D.A., Buck J., Brokstein P., Brottier P.,
 RA Burtis D.C., Butler H., Cadieu E., Cener A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Flischmann W.,
 RA Foster C., Gabriellan A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lascko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Munzy D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclebo J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195 (2000).
 [2]

RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champagne M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Murthy D.M., Nelson C.R.,
 RA Paclebo J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svitskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
 RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
 RT *melanogaster* euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 [3]

RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svitskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.,
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RT a genomic perspective.";
 RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
 [4]

RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter U., Russo S., Schroeder A.U., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.,
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 [5]

RP NUCLEOTIDE SEQUENCE.
 RG Berkeley *Drosophila* Genome Project;
 RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
 RA Hoskins R., Stapleton M., Paclebo J., Park S., Svitskas R., Smith E.,
 RA Yu C., Rubin G.,

RT "Drosophila melanogaster release 4 sequence.";
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RG Flybase;
 RL Submitted (MAR-2005) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AE003435; AAFA6107.1; -, Genomic DNA.
 DR EMBL; CG4790; Drosophila melanogaster.
 DR Flybase; FBgn0005390; CG4790.
 DR Flybase; FBgn0005390; fs(1)M3.
 SQ SEQUENCE 1793 AA; 202233 MW; DDC38A764E301CD CRC64;

Query Match 58.1%; Score 18; DB 2; Length 1793;
 Best Local Similarity 22.2%; Pred. No. 73;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXKC 12
 DB 291 CTSSASASC 299

RESULT 26
 O8MP07 DROME PRELIMINARY; PRT; 1836 AA.
 ID O8MP07 DROME PRELIMINARY;
 AC O8MP07

DT 01-OCT-2002 (TRENBLREL. 22, Created)
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE Poliole precursor.
 GN Name=fs(1)M3; Synonyms=fs(1)ph;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;

RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21956567; PubMed=11959640; DOI=10.1101/gad.223902;
 RA Jimenez G., Gonzalez-Reyes A., Casanova U.,
 RT "Cell surface proteolins Nasrat and Poliole stabilize the Torso-like
 RT extracellular determinant in *Drosophila* oogenesis.";
 RL Genes Dev. 16:913-918 (2002).
 DR EMBL; AJ487522; CAD31790.1; -, mRNA.
 DR EMBL; CG4790; Drosophila melanogaster.
 DR Flybase; FBgn0005390; CG4790.
 DR Flybase; FBgn0005390; fs(1)M3.
 FT SIGNAL.
 FT CHAIN 1 25 Potential.
 SQ SEQUENCE 1836 AA; 207449 MW; 4B0FC64DCDBF52BF CRC64;

Query Match 58.1%; Score 18; DB 2; Length 1836;
 Best Local Similarity 22.2%; Pred. No. 74;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXKC 12
 DB 291 CTSSASASC 299

RESULT 27
 Q40515 LEIMA PRELIMINARY; PRT; 1894 AA.
 ID Q40515 LEIMA PRELIMINARY;
 AC Q40515;
 DT 13-SEP-2005 (TRENBLREL. 31, Created)
 DT 13-SEP-2005 (TRENBLREL. 31, Last sequence update)
 DT 13-SEP-2005 (TRENBLREL. 31, Last annotation update)
 DE Calpain-like cysteine peptidase, putative (Cysteine peptidase, clan
 DE ca, family c2, putative).
 GN ORFNames=Lmjf32.0970;
 OS Leishmania major.
 OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 NCBI_TaxID=5664;


```

RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Fiedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrett B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; C7005269; CAJ08617.1; -; Genomic DNA.
SQ SEQUENCE 1894 AA; 199683 MW; 3F71DE2F53BF95BB CRC64;

Query Match
Best Local Similarity 22.2%; Score 18; DB 2; Length 1894;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 1785 CAAAAAAC 1793

RESULT 28
Q4S488.TETNG PRELIMINARY; PRT; 2087 AA.
AC Q4S488;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAP14743, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00024299001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxId=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Jutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Bernardis V.,
RA Crnaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McKernan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Winkler P., Schachter V., Queller F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollins H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG GenomeScope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
DR EMBL; CAE01014743; CAG04544.1; -; Genomic DNA.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_CA_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR00173; GAP_dhdtrogenase.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR012680; Laminin_G_2.
DR Pfam; PF00028; Cadherin; 8.
DR Pfam; PF01049; Cadherin_C; 1.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF02210; Laminin_G_2; 2.

```

```

DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 9.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00282; LamG; 2.
DR PROSITE; PS00232; CADHERIN_1; 6.
DR PROSITE; PS00268; CADHERIN_2; 9.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00071; GAPDH; UNKNOWN_1.
DR PROSITE; PS50025; LM_G_DOMAIN; 2.
KW Calcium; EGF-like domain; Transmembrane.
FT NON TER 1
SQ SEQUENCE 2087 AA; 222563 MW; 2C55D0554ECDD7A1 CRC64;

Query Match
Best Local Similarity 22.2%; Score 18; DB 2; Length 2087;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 442 CSSSSAASC 450

RESULT 29
Q9L981.VIRCH PRELIMINARY; PRT; 50 AA.
AC Q9L981;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF1.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxId=666;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=569b;
RA MEDLINE=20476433; PubMed=11021935;
RA Clark C.A., Purins L., Kaewrakon P., Focareta T., Manning P.A.;
RL "The Vibrio cholerae O1 chromosomal integrase.";
RT Microbiology 146:2605-2612(2000).
DR EMBL; AF173596; AAF71194.1; -; Genomic DNA.
SQ SEQUENCE 50 AA; 5564 MW; 1A18FC4D3634E57 CRC64;

Query Match
Best Local Similarity 54.8%; Score 17; DB 2; Length 50;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 28 CASLASSSC 36

RESULT 30
Q4RA29.TETNG PRELIMINARY; PRT; 71 AA.
AC Q4RA29;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAP24619, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00037393001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxId=99883;
RN [1]

```

RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Oquif-Costaz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dosset C., Segurens B.,
RA Anthouard C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Dandlauer V., Jubin C., Castell V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Bernardinis V.,
RA Cruaud C., Duprat S., Brottier P., Goucanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Landet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Croliun H.;
RT "Genome duplication in the teleost fish tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: CAAB01024619; CAG14754.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 71 AA; 8036 MW; 736AF2B349B9AC50 CRC64;
Query Match 54.8%; Score 17; DB 2; Length 71;
Best Local Similarity 22.2%; Pred. No. 66;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CXXXXXXXXC 12
DB 27 CSTSRATC 35
RESULT 31
Q8VD96_MESAU PRELIMINARY; PRT; 96 AA.
ID Q8VD96;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE 10 kDa protein precursor.
GN Name=CC10;
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Cricetinae; Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RA Gutierrez-Sagat R., Nieto A.;
RT "Cloning and sequencing of a cDNA encoding hamster uteroglobin/clara
RT cell 10 kDa protein.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: I37041; AAL31349.1; -; Genomic_DNA.
DR HSSP; P17559; IUTR.
DR SMR; Q8VD96; 20-96.
DR GO: 0005496; F:steroid binding; IEA.
DR InterPro; IPR006039; Utergl.
DR InterPro; IPR003628; Uteroglobin_sub.
DR InterPro; IPR000329; Uteroglobin_subf.
DR InterPro; IPR006038; Uteroglobin_subf.
DR PANTHER; PTHR10136; Uteroglobin_sub; 1.
DR Pfam; PF01099; Uteroglobin; 1.
DR PRINTS; PR00486; UTEROGLBIN.
DR ProDom; PD012475; Uteroglbn_sub; 1.
DR SMART; SM00096; UTG; 1.
DR PROSITE; PS00404; UTEROGLBIN_2; 1.
FT SIGNAL 1 18 Potential.

SQ SEQUENCE 96 AA; 10509 MW; 5EB9CBDD46143389 CRC64;
Query Match 54.8%; Score 17; DB 2; Length 96;
Best Local Similarity 22.2%; Pred. No. 74;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CXXXXXXXXC 12
DB 16 CSSASSDTC 24
RESULT 32
Q4P4H6_USTMA PRELIMINARY; PRT; 108 AA.
ID Q4P4H6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM04987.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantang P., Baldwin J., Barry A.,
RA Bayl T., Blithstein B., Bloom T., Biye J., Boguslavsky L.,
RA Borowsky M., Boukhalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Chestnang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Doolley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Ghitre A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysealis M., Karlsson E.,
RA Kelle C., Kieu A., Kienner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-Toh K., Liu X., Lokysang T., Lokysang Y., Lucien O.,
RA Lui A., Ma L.U., Mabbitt R., MacDonald J., Maclean C., Major J.,
RA Manning J., Marabelli R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., McInee T., Meidim J., Meneus L.,
RA Mesirov J., Mihaliev A., Mihova T., Mikelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okowo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pigani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schubach R., Seaman C., Settillali S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnuez C.,
RA Spencer B., Stalker J., Strange-thomann N., Scavopoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas D., Tchinga P.,
RA Tenzing P., Teffaye S., Theodore J., Thoulungang Y., Topham K.,
RA Towse S., Tsamila T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yedav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AACP0100184; EAK85760.1; -; Genomic_DNA.
DR Hypothetical protein.
SQ SEQUENCE 108 AA; 11407 MW; E641B76DB8C16E7A CRC64;

Query Match 54.8%; Score 17; DB 2; Length 108;
 Best Local Similarity 22.2%; Pred. No. 78;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
 DB 14 CTSANATC 22

RESULT 33

Q9ZSP6 ARATH PRELIMINARY; PRT; 108 AA.

AC Q9ZSP6;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Putative cell wall-plasma membrane disconnecting CLCT protein.
 GN Name=AR1B;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN 1)

RP NUCLEOTIDE SEQUENCE.
 RA Neuhboom L.W., van der Zaal E.J;
 RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.

RESULT 34

Q72D14 DESVH PRELIMINARY; PRT; 115 AA.

AC Q72D14;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedlocusNames=DVU117;
 OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB 8303).
 OC Bacteria; Proteobacteria; DeltaProteobacteria; Desulfovibrionales;
 OC Desulfovibrionaceae; Desulfovibrio.
 OC NCBI_TaxID=882;
 RN 1)

RP NUCLEOTIDE SEQUENCE.
 RA Kolnberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
 RA Kolnberg J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
 RA Daugherty S.C., Deboy R.T., Dodson R.U., Durkin A.S., Madupu R.,
 RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
 RA Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D.,

Query Match 54.8%; Score 17; DB 2; Length 108;
 Best Local Similarity 22.2%; Pred. No. 78;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
 DB 20 CTSATTGTC 28

RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Uterback T.R.,
 RA Feldblum T.V., Wall J.D., Voordouw G., Fraser C.M.;
 RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
 RT Desulfovibrio vulgaris Hildenborough.",
 RL Nat. Biotechnol. 22:554-559(2004).
 DR EMBL, AE017313; AAS95597.1; -, Genomic_DNA.
 DR TIGR, DVU117; -;
 KM Complete proteome; Hypothetical protein.
 SQ SEQUENCE 115 AA; 11838 MW; 8E8A4DDF3EB96E64 CRC64;

Query Match 54.8%; Score 17; DB 2; Length 115;
 Best Local Similarity 22.2%; Pred. No. 80;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
 DB 70 CDTATASC 78

RESULT 35

O6VZRL CNPV PRELIMINARY; PRT; 117 AA.

AC O6VZRL;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE CNP086 TNFR-like protein.
 GN Name=CNP086;
 OS Canaripox virus (CNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Avipoxvirus.
 OC NCBI_TaxID=44088;
 RN 1)

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC VR-111;
 RX PubMed=14671117; DOI=10.1128/JVI.78.1.353-366.2004;
 RA Tulman E.R., Afonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
 RT "The genome of canarypox virus.",
 RL J. Virol. 78:353-366(2004).
 DR EMBL, AY318871; AAR83432.1; -, Genomic_DNA.
 DR GO, GO:0004872; Fireceptor activity; IEA.
 DR InterPro, IPR006209; EGF-like.
 DR InterPro, IPR001368; TNFR_c6.
 DR Pfam, PF00020; TNFR_c6; 1-
 DR SMART, SM00208; TNFR; 2-
 DR PROSITE, PS01186; EGF 2; UNKNOWN_1
 DR PROSITE, PS00652; TNFR_NGFR_1; UNKNOWN_1.
 DR PROSITE, PS50050; TNFR_NGFR_2; 1.
 DR PROSITE, PS50050; TNFR_NGFR_2; 1.
 SQ SEQUENCE 117 AA; 13113 MW; 06E99CAB6A41135A CRC64;

Query Match 54.8%; Score 17; DB 2; Length 117;
 Best Local Similarity 22.2%; Pred. No. 80;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
 DB 50 CTSSTTTC 58

RESULT 36

Q4TWL2 ASF PRELIMINARY; PRT; 119 AA.

AC Q4TWL2;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE 9RL protein (Fragment).
 GN Name=B602L;
 OS African swine fever virus (ASFV).
 OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfarvirus.
 OC NCBI_TaxID=10497;
 RN 1)

RP NUCLEOTIDE SEQUENCE.
 RA Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D.,

```

RC STRAIN=Portalegre 90;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
fever virus.";
RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028332; AAY46118.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 54.8%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 33 CASTGASTC 41

RESULT 37
Q4TWL3 ASF PRELIMINARY; PRT; 119 AA.
AC Q4TWL3.
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Iisbon 60;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
fever virus.";
RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028331; AAY46117.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 54.8%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 33 CASTGASTC 41

RESULT 38
Q4TWL5 ASF PRELIMINARY; PRT; 119 AA.
AC Q4TWL5.
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Coimbra 87;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
fever virus.";
RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028329; AAY46115.1; -; Genomic_DNA.
FT NON_TER 1

```

```

FT NON_TER 119
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 54.8%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 33 CASTGASTC 41

RESULT 39
Q4TWL6 ASF PRELIMINARY; PRT; 119 AA.
AC Q4TWL6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Matra 86;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
fever virus.";
RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028328; AAY46114.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 119 AA; 11863 MW; 3A062A08CBA8D31F CRC64;

Query Match
Best Local Similarity 54.8%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 33 CASTGASTC 41

RESULT 40
Q4TWL7 ASF PRELIMINARY; PRT; 119 AA.
AC Q4TWL7.
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Barrancos 93;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
fever virus.";
RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028327; AAY46113.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 54.8%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 4 CXXXXXXC 12
DB 33 CASTGASTC 41

RESULT 41

Q4TWL8 ASF PRELIMINARY; PRT; 119 AA.
AC Q4TWL8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Almodovar 99;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine fever virus."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028326; AAY46112.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match 54.8%; Score 17; DB 2; Length 119;
Best Local Similarity 22.2%; Pred. No. 81;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 33 CASTGASTC 41

RESULT 42

Q4TWL9 ASF PRELIMINARY; PRT; 119 AA.
AC Q4TWL9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Almodovar 99/NE1;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine fever virus."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028325; AAY46111.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match 54.8%; Score 17; DB 2; Length 119;
Best Local Similarity 22.2%; Pred. No. 81;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 33 CASTGASTC 41

RESULT 43
Q4TWM0 ASF

ID Q4TWM0 ASF PRELIMINARY; PRT; 119 AA.

AC Q4TWM0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Almodovar 99/E2;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine fever virus."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028324; AAY46110.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match 54.8%; Score 17; DB 2; Length 119;
Best Local Similarity 22.2%; Pred. No. 81;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 33 CASTGASTC 41

RESULT 44

Q5YPN5 NOCFA PRELIMINARY; PRT; 125 AA.
AC Q5YPN5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=nfa50040;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K., Shiba T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IFM 10152."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR EMBL; AP006618; BAD59856.1; -; Genomic_DNA.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 125 AA; 12749 MW; 1FD69A1CADF924EB CRC64;

Query Match 54.8%; Score 17; DB 2; Length 125;
Best Local Similarity 22.2%; Pred. No. 83;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 96 CRSSATTC 104

RESULT 45

Q4TZU4 TETNG PRELIMINARY; PRT; 133 AA.
AC Q4TZU4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 12 SCAF10174, whole genome shotgun sequence.

```

DE (Fragment)
GN ORFNames=GSTENG00008171001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCB1_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Coataz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Antouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bismont C., Skali Z., Cattolico L., Poulain J., De Berardinis V.,
RA Craud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Queller F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAEO1010174; CAF92788.1; -; Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 133 AA; 14711 MW; 528890284B8084F4 CRC64;

Query Match 54.8%; Score 17; DB 2; Length 133;
Best Local Similarity 22.2%; Pred. No. 85;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXC 12
Db 122 CQATSTTC 130

RESULT 46
O6ZRF5 HUMAN
ID O6ZRF5 HUMAN PRELIMINARY; PRT; 134 AA.
AC O6ZRF5;
DT 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLN46388.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OX NCB1_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Thymus.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saiko K., Nishikawa T., Kimura K.,
RA Yamashta H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Magatsuna M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Naghatai K.,
RA Masuno Y., Nagai K., Isogai T.,
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128253; BAC87355.1; -; mRNA.
SQ SEQUENCE 134 AA; 14481 MW; 03BA7E349C87BC6A CRC64;

```

```

Query Match 54.8%; Score 17; DB 2; Length 134;
Best Local Similarity 22.2%; Pred. No. 85;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXC 12
Db 27 CGASAAAAC 35

RESULT 47
O9CMG3 MOUSE
ID O9CMG3 MOUSE PRELIMINARY; PRT; 147 AA.
AC O9CMG3;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE Mus musculus ES cells cDNA, RIKEN full-length enriched library,
DE clone:2410101B22 product:LENG5 PROTEIN homolog.
GN Name=Leng5;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=C57BL/6J;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=C57BL/6J;
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai T., Shinagawa A., Shibata K., Yoshino M., Itoh Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saiko R.,
RA Kadota K., Matuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Batsch G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.U., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=C57BL/6J;
RC The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=C57BL/6J;
RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayata N., Sugihara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalization and subcloning of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RA MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Mishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanishi M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-Format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arikawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK010759; BAB27164.1; -; mRNA.
DR Ensembl; ENSMUSG0000035585; Mus musculus.
DR MGI; MGI:1913328; Leng5.
SQ SEQUENCE 147 AA; 15856 MW; 988434686F9EA74 CRC64;
Query Match 54.8%; Score 17; DB 2; Length 147;
Best Local Similarity 22.2%; Pred. No. 88;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CXXXXXXC 12
Db 19 CATACTCTC 27

RESULT 48
O6F340_ORYSA PRELIMINARY; PRT; 155 AA.
AC O6F340;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein OSUNBA0088106.14.
GN Name=OSUNBA0088106.14;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_Taxid=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Chang C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC129718; AAT69648.1; -; Genomic_DNA.
DR Gramene; O6F340; -;
KW Hypothetical protein.
SQ SEQUENCE 155 AA; 16979 MW; 0EADP4C906C39295 CRC64;
Query Match 54.8%; Score 17; DB 2; Length 155;
Best Local Similarity 22.2%; Pred. No. 90;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
Db 146 CRASASTC 154

RESULT 49
O4TWL4_ASF PRELIMINARY; PRT; 155 AA.
AC O4TWL4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfarivirus.
OX NCBI_Taxid=10497;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Duarte M.M., Cruz M.B.;
RA "B602L (9RL) partial gene of Portuguese isolates of African swine
RT fever virus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028330; AAY46116.1; -; Genomic_DNA.
FT NON_TER 1 155
FT NON_TER 1 155
SQ SEQUENCE 155 AA; 15125 MW; 31AD66598783CE06 CRC64;
Query Match 54.8%; Score 17; DB 2; Length 155;
Best Local Similarity 22.2%; Pred. No. 90;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CXXXXXXC 12
Db 69 CASTGASTC 77

RESULT 50
CI025_HUMAN STANDARD; PRT; 157 AA.
AC CI025_HUMAN
AC O81W50; Q5T591; Q8TA28;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Protein C9orf25.
GN Name=C9orf25;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;
OC Homo.
OX NCBI_Taxid=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15164053; DOI=10.1038/nature02465;
RA Humphrey S.J., Oliver K., Hunt A.R., Plumb R.W., Loveland J.E.,
RA Howe K.L., Andrews T.D., Searle S., Hunt S.E., Scott C.E., Jones M.C.,
RA Ainscough R., Almeida J.P., Ambrose K.D., Ashwell R.I.S.,
RA Babage A.K., Babage S., Baguley C.L., Bailey J., Banerjee R.,
RA Barker D.J., Barlow K.F., Bates K., Beasley H., Beasley O., Bird C.P.,
RA Bray-Allen S., Brown A.J., Brown J.Y., Burford D., Burrill W.,
RA Burton J., Carter C., Carter N.P., Chapman J.C., Chen Y., Clarke G.,
RA Clark S.Y., Clee C.M., Clegg S., Collier R.E., Corby N., Croxier M.,
RA Cumming A.T., Davies J., Dhami P., Dunn M., Dutta I., Dyer L.W.,
RA Earlrowl M.E., Faulkner L., Fleming C.J., Franchi A.,
RA Frankland J.A., French L., Fricker D.G., Garner P., Garnett J.,
RA Ghori J., Gilbert J.G.R., Glison C., Graham D.V., Gribble S.,
RA Griffiths C., Griffiths-Jones S., Grocock R., Guy J., Hall R.E.,
RA Hammond S., Harley J.L., Harrison E.S.I., Hart E.A., Heath P.D.,
RA Henderson C.D., Hopkins B.L., Howard P.J., Howden P.J., Huckle E.,
RA Johnson C., Johnson D., Joy A.A., Kay M., Keenan S., Kershaw J.K.,
RA Kimberley A.M., King A., Knights A., Laird G.K., Langford C.,

RA Lawlor S., Leongamornlert D.A., Leverisha M., Lloyd C., Lloyd D.M.,
RA Lovell J., Martin S., Mashreghi-Mohammadi M., Matthews L., McLaren S.,
RA McLay K.E., McMurray A., Milne S., Nickerson T., Nisbett J.,
RA Nordiek G., Pearce A.V., Peck A.I., Porter K.M., Pandian R.,
RA Pelan S., Phillimore B., Povey S.K., Ramsey Y., Rand V., Scharfe M.,
RA Sehra H.K., Showmken R., Sims S.K., Stuce C.D., Smith M.,
RA Steward C.A., Swarbreck D., Symore N., Tester J., Thorpe A.,
RA Tracey A., Tromans A., Thomas D.W., Wall M., Wallis J.M., West A.P.,
RA Whitehead S.L., Willey D.L., Williams S.A., Wilming L., Wray P.W.,
RA Young L., Ashurst J.L., Coulson A., Blocker H., Durbin R.,
RA Sulston J.E., Hubbard T., Jackson M.J., Bentley D.R., Beck S.,
RA Rogers J., Dunham I.,
RT "DNA sequence and analysis of human chromosome 9.",
RL Nature 429:369-374 (2004).
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullan S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences",
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AL160270; CA113153.1; ALT INIT; Genomic DNA.
DR EMBL; AL356494; CA113153.1; JOINED; Genomic DNA.
DR EMBL; AL356494; CA115974.1; ALT INIT; Genomic DNA.
DR EMBL; AL160270; CA115974.1; JOINED; Genomic DNA.
DR EMBL; BC025348; AAH25348.1; -; mRNA.
DR EMBL; BC041009; AAH41009.1; -; mRNA.
DR Ensembl; ENSG00000164970; Homo sapiens.
DR HGN; HGNC:19920; C9orf25.
FT CONFLICT 106 106 Missing (in Ref. 2; AAH41009).
SQ SEQUENCE 157 AA; 17365 MW; 577A4FAA6B7DC74 CRC64;

Query Match 54.8%; Score 17; DB 1; Length 157;
Best Local Similarity 22.2%; Pred. No. 90;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12
DB 146 COATSTAC 154

Search completed: January 4, 2006, 16:08:57
Job time : 127.957 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 ; Search time 109.496 Seconds
(without alignments)
64.204 Million cells updates/sec

Title: US-09-932-322-4

Perfect score: 32

Sequence: 1 XXXXXXXXXXXXXXX 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

1: A.Geneseq.21.*
2: geneseqp1980s.*
3: geneseqp1990s.*
4: geneseqp2000s.*
5: geneseqp2001s.*
6: geneseqp2002s.*
7: geneseqp2003as.*
8: geneseqp2003bs.*
9: geneseqp2004s.*
10: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	56.2	30	AD247471	Ad247471 Aliu sense
2	18	56.2	89	ADP30533	Adp30533 Human sec
3	18	56.2	89	ADP30531	Adp30531 Human sec
4	18	56.2	109	AAW93527	Aaw93527 Human pol
5	18	56.2	109	ADL31230	Adl31230 Human pro
6	18	56.2	109	ADP30690	Adp30690 Human sec
7	18	56.2	144	ADP31474	Adp31474 Human sec
8	18	56.2	167	ABO69151	AbO69151 Pseudomon
9	18	56.2	169	ADL31285	AdL31285 Human the
10	18	56.2	171	ADP30794	Adp30794 Human sec
11	18	56.2	171	ADP30793	Adp30793 Human sec
12	18	56.2	183	ADP30806	Adp30806 Human sec
13	18	56.2	191	ABO76619	AbO76619 Pseudomon
14	18	56.2	214	ADZ56213	Adz56213 Human K1A
15	18	56.2	228	ADP30921	Adp30921 Human sec
16	18	56.2	252	ADP31485	Adp31485 Human sec
17	18	56.2	258	ADP30479	Adp30479 Human sec
18	18	56.2	264	ADP31412	Adp31412 Human sec
19	18	56.2	270	ADP30500	Adp30500 Human sec
20	18	56.2	270	ADP31321	Adp31321 Human sec
21	18	56.2	294	ADP31473	Adp31473 Human sec
22	18	56.2	297	ADP31192	Adp31192 Human sec
23	18	56.2	328	AAW25285	Aaw25285 Human pro
24	18	56.2	339	ADP30702	Adp30702 Human sec

25	18	56.2	348	ADP31441	Adp31441 Human sec
26	18	56.2	356	ABG21039	Abg21039 Novel hum
27	18	56.2	357	ADP31267	Adp31267 Human sec
28	18	56.2	357	ADP30505	Adp30505 Human sec
29	18	56.2	360	ADP31439	Adp31439 Human sec
30	18	56.2	390	ADP31218	Adp31218 Human sec
31	18	56.2	421	ADP31159	Adp31159 Human sec
32	18	56.2	426	ADP31495	Adp31495 Human sec
33	18	56.2	438	ABO82561	AbO82561 Pseudomon
34	18	56.2	453	ADP31465	Adp31465 Human sec
35	18	56.2	471	ADP30854	Adp30854 Human sec
36	18	56.2	525	ADP31227	Adp31227 Human sec
37	18	56.2	549	ADP30855	Adp30855 Human sec
38	18	56.2	555	ADP31416	Adp31416 Human sec
39	18	56.2	555	ADP31417	Adp31417 Human sec
40	18	56.2	588	ADP31699	Adp31699 Human sec
41	18	56.2	591	ADZ56209	Adz56209 Human K1A
42	18	56.2	591	ADZ56223	Adz56223 Human K1A
43	18	56.2	600	ADP30865	Adp30865 Human sec
44	18	56.2	604	ADP30940	Adp30940 Human sec
45	18	56.2	604	ADP30941	Adp30941 Human sec
46	18	56.2	605	ADP30507	Adp30507 Human sec
47	18	56.2	617	ADP31657	Adp31657 Human sec
48	18	56.2	626	ABR41596	AbR41596 Human DIT
49	18	56.2	645	ADP30858	Adp30858 Human sec
50	18	56.2	653	AAy94907	Aay94907 Human sec
51	18	56.2	653	ADZ73254	Adz73254 Human kid
52	18	56.2	653	ADW42108	Adw42108 Protein e
53	18	56.2	653	ADH9077	Adh9077 TAT290. 1
54	18	56.2	653	ADL10948	AdL10948 Human the
55	18	56.2	653	ADZ56204	Adz56204 Human K1A
56	18	56.2	653	ADZ56214	Adz56214 Human K1A
57	18	56.2	660	ADP31677	Adp31677 Human sec
58	18	56.2	665	ADP30571	Adp30571 Human sec
59	18	56.2	681	ADP31582	Adp31582 Human sec
60	18	56.2	690	ADP30750	Adp30750 Human sec
61	18	56.2	708	ADP31191	Adp31191 Human sec
62	18	56.2	711	ADP31440	Adp31440 Human sec
63	18	56.2	728	ADP30508	Adp30508 Human sec
64	18	56.2	739	ADP31196	Adp31196 Human sec
65	18	56.2	750	ADP30576	Adp30576 Human sec
66	18	56.2	757	ADP30925	Adp30925 Human sec
67	18	56.2	762	ADP30891	Adp30891 Human sec
68	18	56.2	771	ADP31469	Adp31469 Human sec
69	18	56.2	772	ADP30936	Adp30936 Human sec
70	18	56.2	774	ADP30506	Adp30506 Human sec
71	18	56.2	774	ADP30780	Adp30780 Human sec
72	18	56.2	795	ADP30781	Adp30781 Human sec
73	18	56.2	803	ADP31332	Adp31332 Human sec
74	18	56.2	804	ADP31291	Adp31291 Human sec
75	18	56.2	813	ADP30649	Adp30649 Human sec
76	18	56.2	834	ADP30738	Adp30738 Human sec
77	18	56.2	852	ADP30664	Adp30664 Human sec
78	18	56.2	856	ADP30972	Adp30972 Human sec
79	18	56.2	865	ADP30566	Adp30566 Human sec
80	18	56.2	882	ADP31688	Adp31688 Human sec
81	18	56.2	888	ADP30971	Adp30971 Human sec
82	18	56.2	892	ADP31578	Adp31578 Human sec
83	18	56.2	892	ADP31578	Adp31578 Human sec
84	18	56.2	930	ADP31444	Adp31444 Human sec
85	18	56.2	933	ADP31486	Adp31486 Human sec
86	18	56.2	947	ADP30937	Adp30937 Human sec
87	18	56.2	960	ADP31471	Adp31471 Human sec
88	18	56.2	960	ADP31470	Adp31470 Human sec
89	18	56.2	1017	AAO26719	Aao26719 SR protei
90	18	56.2	1033	ADP30984	Adp30984 Human sec
91	18	56.2	1065	ADP30661	Adp30661 Human sec
92	18	56.2	1065	ADP30973	Adp30973 Human sec
93	18	56.2	1065	ADP31482	Adp31482 Human sec
94	18	56.2	1090	ADQ10187	Adq10187 Human pol
95	18	56.2	1092	ADP31358	Adp31358 Human sec
96	18	56.2	1110	ADP31430	Adp31430 Human sec
97	18	56.2	1116	ADP31692	Adp31692 Human sec

244	17	53.1	68	4	AAm6354	AAm6354 Human bon	317	17	53.1	312	8	ADP31505	Adp31505 Human sec
245	17	53.1	68	4	AAm53966	AAm53966 Human bira	318	17	53.1	312	8	ADP31454	Adp31454 Human sec
246	17	53.1	68	4	ABG48020	ABG48020 Human liv	319	17	53.1	314	7	ADP31452	Adp31452 Pseudomon
247	17	53.1	68	4	AAm01955	AAm01955 Peptide #	320	17	53.1	315	8	ADP31685	Adp31685 Human sec
248	17	53.1	68	5	ABG36002	ABG36002 Human pep	321	17	53.1	318	8	ADP31066	Adp31066 Human sec
249	17	53.1	68	3	AAAB1510	AAAB1510 SRV vlrus	322	17	53.1	313	8	ADP31442	Adp31442 Human sec
250	17	53.1	96	4	AAU65514	AAU65514 Propionib	323	17	53.1	333	8	ADP31239	Adp31239 Human sec
251	17	53.1	96	4	AAU65170	AAU65170 Propionib	324	17	53.1	336	8	ADP30470	Adp30470 Human sec
252	17	53.1	96	6	ABM62033	ABM62033 Propionib	325	17	53.1	337	8	ADP31387	Adp31387 Human sec
253	17	53.1	96	6	ABM62689	ABM62689 Propionib	326	17	53.1	339	8	ADP30519	Adp30519 Human sec
254	17	53.1	104	6	ABM68094	ABM68094 Drosophila	327	17	53.1	345	8	ADP31065	Adp31065 Human sec
255	17	53.1	127	8	ADP31622	ADP31622 Human sec	328	17	53.1	348	8	ADP31374	Adp31374 Human sec
256	17	53.1	139	9	ADP25813	ADP25813 Scortias s	329	17	53.1	349	5	ABM80605	ABM80605 H1rudin p
257	17	53.1	139	9	AEAB06475	AEAB06475 S. sporigi	330	17	53.1	354	8	ADP30783	Adp30783 Human sec
258	17	53.1	139	9	AEAB72826	AEAB72826 Carbohydr	331	17	53.1	354	8	ADP30779	Adp30779 Human sec
259	17	53.1	142	8	ADP31632	ADP31632 Human sec	332	17	53.1	354	8	ADP30784	Adp30784 Human sec
260	17	53.1	149	5	ADK34895	ADK34895 Novel hum	333	17	53.1	354	8	ADP30778	Adp30778 Human sec
261	17	53.1	150	7	ABO82982	ABO82982 Pseudomon	334	17	53.1	357	8	ADP31223	Adp31223 Human sec
262	17	53.1	153	8	ADP30467	ADP30467 Human sec	335	17	53.1	357	8	ADP30581	Adp30581 Human sec
263	17	53.1	153	8	ADP31511	ADP31511 Human sec	336	17	53.1	360	8	ADP31209	Adp31209 Human sec
264	17	53.1	156	7	ABO71866	ABO71866 Pseudomon	337	17	53.1	366	8	ADP31670	Adp31670 Human sec
265	17	53.1	156	8	ADP30759	ADP30759 Human sec	338	17	53.1	369	8	ADP30582	Adp30582 Human sec
266	17	53.1	165	8	ADP31041	ADP31041 Human sec	339	17	53.1	377	8	ADP31689	Adp31689 Human sec
267	17	53.1	168	8	ADP31621	ADP31621 Human sec	340	17	53.1	378	8	ADP31375	Adp31375 Human sec
268	17	53.1	168	8	ADP31099	ADP31099 Human sec	341	17	53.1	381	8	ADP30655	Adp30655 Human sec
269	17	53.1	168	8	ADP30797	ADP30797 Human sec	342	17	53.1	383	8	ADP31094	Adp31094 Human sec
270	17	53.1	168	8	ADP30769	ADP30769 Human sec	343	17	53.1	384	8	ADP30656	Adp30656 Human sec
271	17	53.1	168	8	ADP30771	ADP30771 Human sec	344	17	53.1	390	4	AAU30181	AAU30181 Novel hum
272	17	53.1	169	7	ABO75555	ABO75555 Pseudomon	345	17	53.1	392	8	ADP31061	Adp31061 Human sec
273	17	53.1	169	7	ABO71598	ABO71598 Pseudomon	346	17	53.1	392	8	ADP31055	Adp31055 Human sec
274	17	53.1	172	6	AAU64154	AAU64154 Propionib	347	17	53.1	393	8	ADP31345	Adp31345 Human sec
275	17	53.1	172	6	ABM60673	ABM60673 Propionib	348	17	53.1	399	8	ADP31313	Adp31313 Human sec
276	17	53.1	172	7	ABO74304	ABO74304 Pseudomon	349	17	53.1	399	8	ADP30578	Adp30578 Human sec
277	17	53.1	177	8	ADP31682	ADP31682 Human sec	350	17	53.1	402	8	ADP30742	Adp30742 Human sec
278	17	53.1	180	8	ADP31496	ADP31496 Human sec	351	17	53.1	402	8	ADP31262	Adp31262 Human sec
279	17	53.1	187	3	AAV72094	AAV72094 Human ser	352	17	53.1	411	7	ABO76350	ABO76350 Pseudomon
280	17	53.1	187	5	ABG30791	ABG30791 Human ser	353	17	53.1	414	8	ADP31567	Adp31567 Human sec
281	17	53.1	187	5	AAE17922	AAE17922 Human gen	354	17	53.1	423	8	ADP31323	Adp31323 Human sec
282	17	53.1	189	8	ADP31641	ADP31641 Human sec	355	17	53.1	429	8	ADP30570	Adp30570 Human sec
283	17	53.1	191	7	ABO70585	ABO70585 Pseudomon	356	17	53.1	432	8	ADP31365	Adp31365 Human sec
284	17	53.1	192	8	ADP30573	ADP30573 Human sec	357	17	53.1	442	8	ADP31134	Adp31134 Human sec
285	17	53.1	195	8	ADP30590	ADP30590 Human sec	358	17	53.1	446	6	ABR40850	ABR40850 R1ctinus c
286	17	53.1	200	7	ABO73751	ABO73751 Pseudomon	359	17	53.1	456	8	ADP31224	Adp31224 Human sec
287	17	53.1	202	7	ABM88006	ABM88006 Rice abio	360	17	53.1	471	8	ADP31567	Adp31567 Human sec
288	17	53.1	203	8	ADP30761	ADP30761 Plant ful	361	17	53.1	471	8	ADP30870	Adp30870 Human sec
289	17	53.1	204	7	ABO82303	ABO82303 Pseudomon	362	17	53.1	472	8	ADP31222	Adp31222 Human sec
290	17	53.1	213	8	ADP31348	ADP31348 Human sec	363	17	53.1	474	8	ADP31157	Adp31157 Human sec
291	17	53.1	222	7	ABO7741	ABO7741 Pseudomon	364	17	53.1	478	8	ADP31007	Adp31007 Human sec
292	17	53.1	222	8	ADP31354	ADP31354 Human sec	365	17	53.1	501	8	ADP31689	Adp31689 Human sec
293	17	53.1	225	8	ADP30589	ADP30589 Human sec	366	17	53.1	519	8	ADP31190	Adp31190 Human sec
294	17	53.1	225	8	ADP31531	ADP31531 Human sec	367	17	53.1	522	8	ADP31070	Adp31070 Human sec
295	17	53.1	229	7	ABO80501	ABO80501 Pseudomon	368	17	53.1	528	8	ADP31525	Adp31525 Human sec
296	17	53.1	243	7	ABO77068	ABO77068 Pseudomon	369	17	53.1	531	8	ADP30594	Adp30594 Human sec
297	17	53.1	248	8	ADP30662	ADP30662 Human sec	370	17	53.1	531	8	ADP31696	Adp31696 Human sec
298	17	53.1	249	7	ADP76961	ADP76961 Novel hum	371	17	53.1	543	8	ADP30849	Adp30849 Human sec
299	17	53.1	254	8	ADP30739	ADP30739 Human sec	372	17	53.1	543	8	ADP31506	Adp31506 Human sec
300	17	53.1	262	8	ADP30768	ADP30768 Human sec	373	17	53.1	543	8	ADP31393	Adp31393 Human sec
301	17	53.1	272	8	ADP30693	ADP30693 Human sec	374	17	53.1	548	2	AAW40306	AAW40306 Parapox v
302	17	53.1	273	8	ADP31236	ADP31236 Human sec	375	17	53.1	552	8	ADP31524	Adp31524 Human sec
303	17	53.1	279	8	ADP31509	ADP31509 Human sec	376	17	53.1	559	8	ADP30716	Adp30716 Human sec
304	17	53.1	279	8	ADP31503	ADP31503 Human sec	377	17	53.1	563	8	ADP30881	Adp30881 Human sec
305	17	53.1	279	8	ADP31671	ADP31671 Human sec	378	17	53.1	564	8	ADP30743	Adp30743 Human sec
306	17	53.1	281	7	ABO80453	ABO80453 Pseudomon	379	17	53.1	564	8	ADP31194	Adp31194 Human sec
307	17	53.1	291	8	ADP31243	ADP31243 Human sec	380	17	53.1	573	8	ADP31078	Adp31078 Human sec
308	17	53.1	291	8	ADP31242	ADP31242 Human sec	381	17	53.1	575	8	ADP31133	Adp31133 Human sec
309	17	53.1	294	8	ADP31076	ADP31076 Human sec	382	17	53.1	585	8	ADP31392	Adp31392 Human sec
310	17	53.1	294	8	ADP31045	ADP31045 Human sec	383	17	53.1	588	8	ADP31458	Adp31458 Human sec
311	17	53.1	297	8	ADP31312	ADP31312 Human sec	384	17	53.1	588	8	ADP30877	Adp30877 Human sec
312	17	53.1	300	8	ADP31301	ADP31301 Human sec	385	17	53.1	588	8	ADP31680	Adp31680 Human sec
313	17	53.1	302	8	ADP31030	ADP31030 Human sec	386	17	53.1	603	8	ADP31150	Adp31150 Human sec
314	17	53.1	302	8	ADP31032	ADP31032 Human sec	387	17	53.1	612	8	ADP31169	Adp31169 Human sec
315	17	53.1	302	8	ADP31025	ADP31025 Human sec	388	17	53.1	618	8	ADP31467	Adp31467 Human sec
316	17	53.1	306	8	ADP31205	ADP31205 Human sec	389	17	53.1	618	8	ADP31466	Adp31466 Human sec


```
974 16 50.0 366 8 ADP31075 Adp31075 Human sec
975 16 50.0 369 7 ABO72044 Abo72044 Pseudomon
976 16 50.0 369 8 ADP31289 Adp31289 Human sec
977 16 50.0 372 8 ADP31610 Adp31610 Human sec
978 16 50.0 375 9 ADP06695 Adp06695 Cyclin-de
979 16 50.0 380 3 AAG04919 Aag04919 Arabidops
980 16 50.0 380 3 AAG32553 Aag32553 Arabidops
981 16 50.0 381 2 AAY38570 Aay38570 Neisseria
982 16 50.0 381 2 AAY75046 Aay75046 Neisseria
983 16 50.0 381 2 AEB48989 Aeb48989 N. mening
984 16 50.0 382 2 AAY38571 Aay38571 Neisseria
985 16 50.0 382 3 AAY75044 Aay75044 Neisseria
986 16 50.0 382 4 ABB70181 ABB70181 Drosophil
987 16 50.0 382 9 AEB48990 Aeb48990 N. gonorr
988 16 50.0 384 8 ADP31338 Adp31338 Human sec
989 16 50.0 384 8 ADP31318 Adp31318 Human sec
990 16 50.0 387 8 ADP31158 Adp31158 Human sec
991 16 50.0 387 9 ADP47819 Adp47819 Amino aci
992 16 50.0 389 8 ADX78458 Adx78458 Plant ful
993 16 50.0 389 8 ADY24278 Ady24278 Plant ful
994 16 50.0 391 7 ADN31030 Adn31030 Fruit fly
995 16 50.0 396 8 ADP31625 Adp31625 Human sec
996 16 50.0 396 8 ADP31378 Adp31378 Human sec
997 16 50.0 398 4 AAM40928 Aam40928 Human con
998 16 50.0 398 7 ADF60215 Adf60215 Human con
999 16 50.0 398 7 ABO71161 Abo71161 Pseudomon
1000 16 50.0 398 8 ADP31604 Adp31604 Human sec
```

ALIGNMENTS

```
RESULT 1
ID AD247471 standard; peptide; 30 AA.
XX AD247471:
AC
XX 30-JUN-2005 (first entry)
XX
DE Alu sense peptide SEQ ID NO 50.
XX
KM hyaluronan synthase; HAS; hyaluronidase; HYAL; HAS1; HAS2; HAS3; HYAL1;
KM HYAL2; HAL3; PH-20; cancer; hyperproliferation; inflammation; anabolic;
KM hypertensive; gynaecological; neuroprotective; anti-neuritic; cytoskeletal;
KM anti-inflammatory; endocrine-gen; immunosuppressive; gene therapy.
XX
OS Unidentified.
XX
PN WO2005035548-A1.
XX
PD 21-APR-2005.
XX
PF 11-OCT-2004; 2004WO-AU001383.
XX
PR 10-OCT-2003; 2003AU-00905551.
PR 01-DEC-2003; 2003AU-00906658.
XX
PA (MEDI-) MEDITECH RES LTD.
XX
PI Brown TJ, Brownlee GR;
XX
DR WPI; 2005-315540/32.
XX
PT Novel compound capable of reducing level, function or activity of
PT hyaluronan synthase or hyaluronidase, useful for treating cancer,
PT hyperproliferative condition, A-beta-lipoproteinemia, A-V, A beta-2-
PT microglobulin amyloidosis.
XX
PS Example 22; SEQ ID NO 50; 277bp; English.
XX
CC This invention describes a novel nucleic acid molecule capable of
CC reducing the level of hyaluronan synthase (HAS) or hyaluronidase (HYAL)
```

or the function or activity of HAS or HYAL. HAS is selected from HAS1, HAS2 and HAS3. HYAL is chosen from HYAL1, HYAL2, HAL3 and PH-20. The nucleic acid molecule is an oligonucleotide or its chemically modified form comprising a chemically modified backbone or a non-natural internucleoside linkage. The compound is an interactive molecule capable of binding or otherwise associating with HAS and/or HYAL to reduce HAS and/or HYAL function or activity. The novel molecule is useful for treating cancer, hyperproliferative conditions or inflammatory conditions. The products of the invention are useful for treating or prophylaxis of a condition in subject, which involves administering to the subject, an HAS and/or HYAL level-reducing or HAS and/or HYAL activity reducing effective amount of active molecule. The products of the invention have anabolic, hypertensive, gynaecological, neuroprotective, anti-neuritic, cytoskeletal, anti-inflammatory, endocrine-gen, and immunosuppressive activity. This sequence represents a peptide used in the method of the invention. NOTE: The specification describes SEQ ID NO 27-51 as nucleotide primers however the Sequence Listing represents the nucleotides in a three letter amino acid code. The nucleotide sequences have been made and are represented in { }.

Sequence 30 AA;

Query Match 56.2%; Score 18; DB 9; Length 30;
Best Local Similarity 20.0%; Pred. No. 2.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 13
Db 18 CTTAAATATAC 27

```
RESULT 2
ADP30533
ID ADP30533 standard; protein; 89 AA.
XX
AC ADP30533;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1300.
XX
KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
```


17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-047136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX
DR WPI; 2004-348438/32.
XX
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 2531; 428bp; English.
XX
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
XX
SQ Sequence 89 AA;

Query Match 56.2%; Score 18; DB 8; Length 89;
Best Local Similarity 20.0%; Pred. No. 3.1;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXX 13
Db 24 CTTAAAAAAC 33
RESULT 3
ADP30531
ID ADP30531 standard; protein; 89 AA.
XX
XX
AC ADP30531;
XX
DT 12-AUG-2004 (first entry)
XX
XX
DE Human secreted protein SEQ ID #1298.
XX
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
XX
28-AUG-2003; 2003WO-US026780.
XX
XX
29-AUG-2002; 2002US-0406576P.
XX
XX
29-AUG-2002; 2002US-0406579P.
XX
XX
29-AUG-2002; 2002US-0406585P.
XX
XX
29-AUG-2002; 2002US-0406588P.
XX
XX
29-AUG-2002; 2002US-0406608P.
XX
XX
29-AUG-2002; 2002US-0406611P.
XX
XX
29-AUG-2002; 2002US-0406612P.
XX
XX
29-AUG-2002; 2002US-0406616P.
XX
XX
29-AUG-2002; 2002US-0406640P.
XX
XX
29-AUG-2002; 2002US-0406642P.
XX
XX
29-AUG-2002; 2002US-0406646P.
XX
XX
29-AUG-2002; 2002US-0406653P.
XX
XX
29-AUG-2002; 2002US-0406655P.
XX
XX
29-AUG-2002; 2002US-0406666P.
XX
XX
17-SEP-2002; 2002US-0410947P.
XX
XX
17-SEP-2002; 2002US-0410948P.
XX
XX
17-SEP-2002; 2002US-0410949P.
XX
XX
17-SEP-2002; 2002US-0410953P.
XX
XX
17-SEP-2002; 2002US-0410957P.
XX
XX
17-SEP-2002; 2002US-0410958P.
XX
XX
17-SEP-2002; 2002US-0410960P.
XX
XX
17-SEP-2002; 2002US-0410961P.
XX
XX
17-SEP-2002; 2002US-0410962P.
XX
XX
17-SEP-2002; 2002US-0411019P.
XX
XX
17-SEP-2002; 2002US-0411022P.
XX
XX
17-SEP-2002; 2002US-0411023P.
XX
XX
17-SEP-2002; 2002US-0411024P.
XX
XX
17-SEP-2002; 2002US-0411032P.
XX
XX
17-SEP-2002; 2002US-0411035P.
XX
XX
17-SEP-2002; 2002US-0411037P.
XX
XX
17-SEP-2002; 2002US-0411041P.
XX
XX
17-SEP-2002; 2002US-0411045P.
XX
XX
17-SEP-2002; 2002US-0411046P.
XX
XX
17-SEP-2002; 2002US-0411048P.
XX
XX
17-SEP-2002; 2002US-0411052P.
XX
XX
17-SEP-2002; 2002US-0411055P.
XX
XX
17-SEP-2002; 2002US-0411073P.
XX
XX
17-SEP-2002; 2002US-0411082P.
XX
XX
17-SEP-2002; 2002US-0411101P.
XX
XX
17-SEP-2002; 2002US-0411111P.
XX
XX
18-APR-2003; 2003US-0463700P.

```

PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 14-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2529; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPOWEB and is not in the specification.
XX
XX SQ Sequence 89 AA;

Query Match          56.2%; Score 18; DB 8; Length 89;
Best Local Similarity 20.0%; Pred. No. 3.1;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 24 CTTAAAAAAC 33

RESULT 4
ID AAM93527
ID AAM93527 standard; protein; 109 AA.
XX
XX AAM93527;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polypeptide, SEQ ID NO: 3263.
XX
XX Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
XX Homo sapiens.

```

```

XX
XX EP1130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-00114089.
XX
XX 08-JUL-1999; 99JP-00194486.
XX 11-JAN-2000; 2000JP-00118774.
XX 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Iehi S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
DR WPI; 2001-524255/58.
XX
XX N-PSDB; AAK94457.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their use
XX in genetic manipulation.
XX
XX Claim 8; SEQ ID NO 3263; 1380bp + Sequence listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been isolated
XX and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX been determined. Primers for synthesizing the full length cDNA are useful
XX for clarifying the function of the protein encoded by the cDNA. The full
XX length clones were obtained by construction of full length enriched cDNA
XX libraries that were synthesised by the oligo-capping method. The primers
XX enable the production of the full length cDNA easily without any special
XX methods. The present sequence is a polypeptide encoded by a full length
XX human cDNA of the invention. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in CD-ROM
XX format directly from EPO

XX SQ Sequence 109 AA;

Query Match          56.2%; Score 18; DB 4; Length 109;
Best Local Similarity 20.0%; Pred. No. 3.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 4 CSSSATSTSC 13

RESULT 5
ID ADL31230
ID ADL31230 standard; protein; 109 AA.
XX
XX ADL31230;
XX
XX 20-MAY-2004 (first entry)
XX
XX Human protein encoded by a full length cDNA clone SegID 3263.
XX
XX human; medicine; signal transduction; glycoprotein; transcription;
XX oligo-capping method.
XX
XX Homo sapiens.
XX
XX EP1396543-A2.
XX
XX 10-MAR-2004.
XX
XX 07-JUL-2000; 2003EP-00025638.
XX
XX 08-JUL-1999; 99JP-00194486.
XX 11-JAN-2000; 2000JP-00118774.
XX 02-MAY-2000; 2000JP-00183865.
XX
XX 07-JUL-2000; 2000EP-00114089.

```

XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H,
XX WPI; 2004-204755/20.
DR N-PSDB; ADL31229.
XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT length human cDNAs.
XX Example 1; SEQ ID NO 3263; 1340bp; English.
XX This invention relates to a novel primers useful for synthesizing full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polypeptide sequence is a full
CC length human protein of the invention.
XX Sequence 109 AA;
SQ
Query Match 56.2%; Score 18; DB 8; Length 109;
Best Local Similarity 20.0%; Pred. No. 3.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXC 13
Db 4 CXXXXXXXC 13
ADP30690 standard; protein; 109 AA.
ID ADP30690 standard; protein; 109 AA.
XX ADP30690;
AC
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1457.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
XX WO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PF
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485233P.
PR 08-JUL-2003; 2003US-0485244P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486896P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;
PI Pierre K, Wang Y, Wong JCB, Wu G, Zhang H;
XX WPI; 2004-348438/32.
DR
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2688; 428bp; English.
PS
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is

CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 109 AA;
Query Match 56.2%; Score 18; DB 8; Length 109;
Best Local Similarity 20.0%; Pred. No. 3.2;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 CXXXXXXXXC 13
|
Db 96 CTTTATATC 105
RESULT 7
ADP31474
ID ADP31474 standard; protein; 144 AA.
XX
AC ADP31474;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2241.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
PN WQ02004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411012P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485235P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493373P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kotakota S, Haisan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 3472; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 144 AA;
Query Match 56.2%; Score 18; DB 8; Length 144;
Best Local Similarity 20.0%; Pred. No. 3.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXC 13
|
Db 72 CAATAAATC 81
RESULT 8
ABO69151
ID ABO69151 standard; protein; 167 AA.
XX
AC ABO69151;
XX
XX
XX 29-JUL-2004 (first entry)
DT
XX
DE Pseudomonas aeruginosa polypeptide #1326.

```

XX  Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX  Pseudomonas aeruginosa.
OS
XX  US6551795-B1.
XX
XX  22-APR-2003.
XX
XX  18-FEB-1999; 99US-00252991.
XX
XX  18-FEB-1998; 98US-0074788P.
XX
XX  27-JUL-1998; 98US-0094190P.
XX
XX  (GENO-) GENOME THERAPEUTICS CORP.
XX
XX  Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI  WPI; 2003-615309/58.
XX
XX  N-PSDB; ABD02722.
XX
XX  Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT  useful as molecular targets for diagnostics, prophylaxis and treatment of
XX  pathological conditions resulting from bacterial infection.
XX
XX  Disclosure; SEQ ID NO 17897; 455bp; English.
XX
XX  The invention relates to Pseudomonas aeruginosa polypeptides and the
XX  polynucleotides encoding them. The sequences are useful in diagnosis and
XX  therapy of pathological conditions, as molecular targets for diagnostics,
XX  prophylaxis and treatment of pathological conditions resulting from a
XX  bacterial infection, for evaluating a compound, such as a polypeptide,
XX  for the ability to bind a P. aeruginosa nucleic acid, as components of
XX  effective antibacterial targets, as targets for antibacterial drugs,
XX  including anti-P. aeruginosa drugs, as templates for recombinant
XX  production of P. aeruginosa-derived peptides or polypeptides, as target
XX  components for diagnosis and/or treatment of P. aeruginosa-caused
XX  infection, and in detection of P. aeruginosa sequences or other sequences
XX  of Pseudomonas species using biochip technology. Sequences AB067826-
XX  AB068336 represent P. aeruginosa polypeptides of the invention. Note: The
XX  sequence data for this patent did not form part of the printed
XX  specification but was obtained in electronic format from USPTO at
XX  seqdata.uspto.gov/sequence.html
XX
XX  Sequence 167 AA;
SQ
XX
XX  Query Match 56.2%; Score 18; DB 7; Length 167;
XX  Best Local Similarity 20.0%; Pred. No. 3.5;
XX  Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY  4 CXXXXXXXXX 13
XX  |
DB  60 CSTSTSTTSC 69
XX

RESULT 9
ADSI2185
ID  ADSI2185 standard; protein; 169 AA.
XX
XX  ADSI2185;
AC
XX
XX  16-DEC-2004 (first entry)
XX
XX  Human therapeutic contig protein - SEQ ID 2422.
XX
XX  antiinflammatory; neuroprotective; antianaemic; cyostatic; vulnary;
XX  inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
XX  aplastic anaemia; cancer; wound healing; gene therapy.
XX
XX  Homo sapiens.
OS
XX
XX  Key Location/Qualifiers
XX  Misc-difference 1. 169

```

```

FT  /label= Unknown, OTHER
FT  /note= "OTHER = In-frame STOP codon"
XX
XX  MO2004080148-A2.
XX
XX  23-SEP-2004.
XX
XX  30-SEP-2003; 2003WO-US030720.
XX
XX  02-OCT-2002; 2002US-0416186P.
XX
XX  (NUVE-) NUVELO INC.
XX
XX  Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;
PI  Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Wang G, Zhou P;
XX  WPI; 2004-668857/65.
XX
XX  N-PSDB; ADS11587.
XX
XX  New polynucleotide, useful in preparing a composition for diagnosing or
PT  treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
XX  aplastic anemia or cancer for promoting wound healing.
XX
XX  Example 2; SEQ ID NO 2422; 718bp; English.
XX
XX  The invention relates to a novel isolated polynucleotide and the encoded
XX  polypeptide. The molecules of the invention demonstrate antiinflammatory,
XX  neuroprotective, antianaemic, cyostatic and vulnary activities and may
XX  be useful in preparing a composition for diagnosing or treating
XX  inflammatory, haematopoietic, immune, neurodegenerative or stem cell
XX  disorders, such as aplastic anaemia or cancer, as well as for promoting
XX  wound healing. The molecules may also be utilised during gene therapy
XX  procedures. The current sequence is that of a human therapeutic contig
XX  protein of the invention.
XX
XX  Sequence 169 AA;
SQ
XX
XX  Query Match 56.2%; Score 18; DB 8; Length 169;
XX  Best Local Similarity 20.0%; Pred. No. 3.5;
XX  Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY  4 CXXXXXXXXX 13
XX  |
DB  81 CAAAAAAMAC 90
XX

RESULT 10
ADP30794
ID  ADP30794 standard; protein; 171 AA.
XX
XX  ADP30794;
AC
XX
XX  12-AUG-2004 (first entry)
XX
XX  Human secreted protein SEQ ID #1561.
XX
XX  Cyostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX  cancer; inflammatory; immune; human secreted protein.
XX
XX  Homo sapiens.
OS
XX
XX  MO2004035732-A2.
XX
XX  29-APR-2004.
XX
XX  28-AUG-2003; 2003WO-US026780.
XX
XX  29-AUG-2002; 2002US-0406576P.
XX
XX  29-AUG-2002; 2002US-0406579P.
XX
XX  29-AUG-2002; 2002US-0406585P.
XX
XX  29-AUG-2002; 2002US-0406588P.
XX
XX  29-AUG-2002; 2002US-0406608P.
XX
XX  29-AUG-2002; 2002US-0406611P.

```

PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-047136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,

PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2792; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic.
CC antiinflammatory, immunosuppressive, antibacterial and vitrucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 171 AA;

Query Match 56.2%; Score 18; DB 8; Length 171;
Best Local Similarity 20.0%; Pred. No. 3.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 13
DB 146 CAATTAATAC 155

RESULT 11
ADP30793
ID ADP30793 standard; protein, 171 AA.
XX
XX ADP30793;
AC
XX
XX 12-AUG-2004 (first entry)
DT
XX
XX Human secreted protein SEQ ID #1560.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Vitrucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
OS
XX
XX WO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PF
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.

PR	19-MAY-2003	2003US-04.71306P
PR	19-MAY-2003	2003US-04.71336P
PR	22-MAY-2003	2003US-04.72420P
PR	22-MAY-2003	2003US-04.72430P
PR	02-JUN-2003	2003US-04.76509P
PR	02-JUN-2003	2003US-04.76560P
PR	09-JUN-2003	2003US-04.76641P
PR	09-JUN-2003	2003US-04.85218P
PR	08-JUL-2003	2003US-04.85223P
PR	08-JUL-2003	2003US-04.85223P
PR	08-JUL-2003	2003US-04.85223P
PR	08-JUL-2003	2003US-04.85223P
PR	14-JUL-2003	2003US-04.86446P
PR	14-JUL-2003	2003US-04.86446P
PR	15-JUL-2003	2003US-04.86591P
PR	15-JUL-2003	2003US-04.86560P
PR	08-AUG-2003	2003US-04.93341P
PR	08-AUG-2003	2003US-04.93341P
PR	08-AUG-2003	2003US-04.93537P
PR	08-AUG-2003	2003US-04.93577P
PR	08-AUG-2003	2003US-04.93577P

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D
PI Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.

SQ Sequence 183 AA;

Query Match	56.2%;	Score 18;	DB 8;	Length 183;
Best Local Similarity	20.0%;	Pred. No. 3.6;		
Matches 2; Conservative	0;	Mismatches 8;	Indels 0;	Gaps 0;

```

QY      4 CXXXXXXXXX 13
         |         |
Db     36 CATTTTTC 45

```

RESULT 13

ID ABO76619 standard; protein; 191 AA.

AC ABO76619;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #8794.

KW Bacterial infection; *Pseudomonas aeruginosa* infection; antibacterial.

Pseudomonas aeruginosa.

PN US6551795-B1.

PD 22-APR-2003.

18-FEB-1999; 99US-00252991.

PR	18-FEB-1998;	98US-0074788P.
PR	27-JUL-1998;	98US-0094190P.
XX		
PA	(GENO-)	GENOME THERAPEUTICS CORP.

Rubinfeld MJ, Nolling J, Deloughery C, Bush D;

DR WPI; 2003-615309/58.

DR N-PSDB; ABD10190.

PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

Disclosure; SEQ ID NO 25365; 455pp; English.

The invention relates to *Pseudomonas aeruginosa* polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a *P. aeruginosa* nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-*P. aeruginosa* drugs, as templates for recombinant production of *P. aeruginosa*-derived peptides or polypeptides, as target components for diagnosis and/or treatment of *P. aeruginosa*-caused infection, and in detection of *P. aeruginosa* sequences or other sequences of *Pseudomonas* species using biotech technology. Sequences AB067826-AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html

Sequence 191 AA;

Query Match	56.2%	Score 18;	DB 7;	Length 191;
Best Local Similarity	20.0%;	Pred. No. 3,6;		
Matches 2; Conservative	0;	Mismatches 8;	Indels 0;	Gaps 0

Qy	4	CXXXXXXXXC	1.3
Db	49	CSTASATSC	58

RESULT 14

ID ADZ56213 standard; protein; 214 AA.

AC ADZ56213;

DT 30-JUN-2005 (first entry)

Human KIAA0779 splice variant clone CLN00149041.a, protein.

KW KIAA0779; gene expression; inflammation; antiinflammatory; cancer;

KW metabolic disorder; metabolic; viral infection; virucide; infection.

OS Homo sapiens.

Homo sapiens.

PN WO2005035569-A2.

PD 21-APR-2005.

12-OCT-2004; 2004WO-US033408.

PR 10-OCT-2003; 2003US-0510612P.

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

Wong JGP, Hestir K, Collins ALT;

DR WPI; 2005-296268/30.

DR N-PSDB; ADZ56200, ADZ56231.
XX New isolated KIAA0779 nucleic acids and polypeptides, useful for
PT diagnosing, preventing and/or treating inflammatory, immune, viral
PT disorders and cancer, such as kidney, lung, ovarian, bladder, breast,
PT prostate and skin cancers.
XX
XX Claim 11; SEQ ID NO 18; 121bp; English.
XX
XX The invention relates to an isolated nucleic acid molecule comprising at
CC least one polynucleotide sequence (appearing as ADZ56196-ADZ56197,
CC ADZ56199-ADZ56200, ADZ56226-ADZ56227 and ADZ56230-ADZ56231), sequences
CC hybridizing to them under high stringency conditions, sequences having at
CC least 80% sequence identity to them, their complements or biologically
CC active fragments. The nucleic acids are splice variants of the human
CC KIAA0779 gene, encoding transmembrane domain protein(s). Also included
CC are a double-stranded isolated nucleic acid molecule comprising the
CC nucleic acid molecule cited above, a vector comprising the isolated
CC expression of the nucleic acid molecule, a recombinant host cell
CC comprising the nucleic acid molecule cited above, an isolated polypeptide
CC (comprising the encoded proteins or the non-transmembrane (TM) regions of
CC the proteins), a method of making a recombinant host cell, a method of
CC making a polypeptide, a method of determining the presence of the nucleic
CC acid molecule cited above in a sample, a method of determining the
CC presence of a specific antibody to the polypeptide of (4) in a sample, a
CC method of determining the presence of the polypeptide in a sample, an
CC antibody specifically binding to and/or interfering with the biological
CC activity of the nucleic acid molecule cited above (or the polypeptide or
CC its biologically active fragment), a composition comprising a
CC pharmaceutical carrier or excipient (and one or more active agents chosen
CC from the nucleic acid molecule cited above, the vector, the polypeptide,
CC and the antibody), a bacteriophage comprising the antibody (or its
CC fragment), a bacterial cell comprising the bacteriophage, a recombinant
CC host cell that produces the antibody, an animal injected with one or more
CC active agents (chosen from the nucleic acid molecule, the vector, the
CC host cell, the polypeptide, and the antibody), a diagnostic kit
CC (comprising a nucleic acid molecule having at least 6 contiguous
CC nucleotides from the nucleic acid molecule cited above, the isolated
CC polypeptide, or antibody, and reagents to carry out an immunoassay), a
CC method of making an antibody, a method of identifying a modulating agent
CC that modulates the biological activity of the polypeptide, a modulator
CC composition comprising a modulator and a pharmaceutical carrier, a method
CC of treating a disease in a subject, and a method of treating cancer,
CC preferably kidney, cervical, squamous lung, ovarian, bladder, breast,
CC endometrial, prostate, and skin cancer, in a subject. The methods and
CC compositions of the present invention are useful for the diagnosis,
CC prevention and/or treatment of inflammatory, immune, viral disorders and
CC cancer, such as kidney, cervical, squamous lung, ovarian, bladder,
CC breast, endometrial, prostate, and skin cancer. The present sequence
CC represents a protein encoded by a cDNA splice variant from the KIAA0779
CC gene.
XX
XX Sequence 214 AA;
SQ
XX
XX Query Match 56 2%; Score 18; DB 9; Length 214;
Best Local Similarity 20.0%; Pred. No. 3.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXX 13
DB 173 CAAAAAAC 182
RESULT 15
ADP30921
XX ADP30921 standard; protein; 228 AA.
XX
XX ADP30921;
XX
XX 12-AUG-2004 (first entry)
DT
XX
DE Human secreted protein SEQ ID #1688.

XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
OS
XX
XX WO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PE
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.

CC The present invention relates to an isolated nucleic acid molecule
 CC encoding a polypeptide which is believed to be cytostatic,
 CC antiinflammatory, immunosuppressive, antibacterial and antiviral. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is
 CC available on WIPOMB and is not in the specification.

XX
 SQ Sequence 252 AA;

Query Match 56.2%; Score 18; DB 8; Length 252;
 Best Local Similarity 20.0%; Pred. No. 3.9;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 210 CTTTATATAC 219

RESULT 17
 ID ADP30479 standard; protein; 258 AA.
 XX
 AC ADP30479;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human secreted protein SEQ ID #1246.
 XX
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
 KM Cancer; inflammatory; immune; human secreted protein.
 OS Homo sapiens.
 XX
 PN WO2004035732-A2.
 XX
 PD 29-APR-2004.
 XX
 PF 28-AUG-2003; 2003WO-US026780.
 XX
 XX 29-AUG-2002; 2002US-0406576P.
 PR 29-AUG-2002; 2002US-0406579P.
 PR 29-AUG-2002; 2002US-0406585P.
 PR 29-AUG-2002; 2002US-0406588P.
 PR 29-AUG-2002; 2002US-0406608P.
 PR 29-AUG-2002; 2002US-0406611P.
 PR 29-AUG-2002; 2002US-0406612P.
 PR 29-AUG-2002; 2002US-0406616P.
 PR 29-AUG-2002; 2002US-0406640P.
 PR 29-AUG-2002; 2002US-0406642P.
 PR 29-AUG-2002; 2002US-0406646P.
 PR 29-AUG-2002; 2002US-0406653P.
 PR 29-AUG-2002; 2002US-0406655P.
 PR 29-AUG-2002; 2002US-0406666P.
 PR 17-SEP-2002; 2002US-0410946P.
 PR 17-SEP-2002; 2002US-0410947P.
 PR 17-SEP-2002; 2002US-0410948P.
 PR 17-SEP-2002; 2002US-0410949P.
 PR 17-SEP-2002; 2002US-0410953P.
 PR 17-SEP-2002; 2002US-0410957P.
 PR 17-SEP-2002; 2002US-0410958P.
 PR 17-SEP-2002; 2002US-0410959P.
 PR 17-SEP-2002; 2002US-0410960P.
 PR 17-SEP-2002; 2002US-0410961P.
 PR 17-SEP-2002; 2002US-0410962P.
 PR 17-SEP-2002; 2002US-0411019P.
 PR 17-SEP-2002; 2002US-0411022P.
 PR 17-SEP-2002; 2002US-0411023P.
 PR 17-SEP-2002; 2002US-0411024P.
 PR 17-SEP-2002; 2002US-0411032P.
 PR 17-SEP-2002; 2002US-0411035P.
 PR 17-SEP-2002; 2002US-0411037P.

PR 17-SEP-2002; 2002US-0411041P.
 PR 17-SEP-2002; 2002US-0411045P.
 PR 17-SEP-2002; 2002US-0411046P.
 PR 17-SEP-2002; 2002US-0411048P.
 PR 17-SEP-2002; 2002US-0411052P.
 PR 17-SEP-2002; 2002US-0411055P.
 PR 17-SEP-2002; 2002US-0411073P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411101P.
 PR 17-SEP-2002; 2002US-0411111P.
 PR 18-APR-2003; 2003US-0463700P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 02-MAY-2003; 2003US-0467230P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 19-MAY-2003; 2003US-0471366P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 22-MAY-2003; 2003US-0472430P.
 PR 09-JUN-2003; 2003US-0476809P.
 PR 09-JUN-2003; 2003US-0476818P.
 PR 08-JUL-2003; 2003US-0485218P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485224P.
 PR 08-JUL-2003; 2003US-0485325P.
 PR 14-JUL-2003; 2003US-0486446P.
 PR 14-JUL-2003; 2003US-0486480P.
 PR 15-JUL-2003; 2003US-0486891P.
 PR 15-JUL-2003; 2003US-0486960P.
 PR 08-AUG-2003; 2003US-0493341P.
 PR 08-AUG-2003; 2003US-0493370P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
 PI Halenbeck RF, Huang MM, Kochakota S, Halshan L, Linnemann T;
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
 DR WPI; 2004-348438/32.
 XX
 XX
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 XX genetic, bacterial and viral diseases.
 XX
 PS Claim 1; SEQ ID NO 2477; 428bp; English.

CC The present invention relates to an isolated nucleic acid molecule
 CC encoding a polypeptide which is believed to be cytostatic,
 CC antiinflammatory, immunosuppressive, antibacterial and antiviral. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is
 CC available on WIPOMB and is not in the specification.

XX
 SQ Sequence 258 AA;

Query Match 56.2%; Score 18; DB 8; Length 258;
 Best Local Similarity 20.0%; Pred. No. 3.9;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 237 CTTTATATAC 246

RESULT 18
 ADP31412

ID ADP31412 standard; protein; 264 AA.
XX
AC ADP31412;
XX
DT 12-AUG-2004 (first entry)
DE Human secreted protein SEQ ID #2179.
XX
XX Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411101P.
XX 17-SEP-2002; 2002US-0411111P.
XX 18-APR-2003; 2003US-0463700P.
XX 18-APR-2003; 2003US-0463708P.
XX 18-APR-2003; 2003US-0463716P.
XX 18-APR-2003; 2003US-0463732P.
XX 02-MAY-2003; 2003US-0467199P.
XX 02-MAY-2003; 2003US-0467201P.
XX 02-MAY-2003; 2003US-0467203P.
XX 02-MAY-2003; 2003US-0467230P.
XX 19-MAY-2003; 2003US-0471306P.
XX 19-MAY-2003; 2003US-0471366P.
XX 22-MAY-2003; 2003US-0472420P.
XX 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493572P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Halahan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1, SEQ ID NO 3410; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic.
XX antinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPOMEB and is not in the specification.
XX
SQ Sequence 264 AA;
XX
Query Match 56.2%; Score 18; DB 8; Length 264;
Best Local Similarity 20.0%; Pred. No. 3.9;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXC 13
Db 186 CATAAATTC 195
RESULT 19
ADP30500
ID ADP30500 standard; protein; 270 AA.
XX
AC ADP30500;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1267.
XX
KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-047136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-048690P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T,
XX Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,
XX
XX WPI; 2004-348438/32.

XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT PT
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 2498; 428pp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 270 AA;
XX
Query Match . 56.2%; Score 18; DB 8; Length 270;
Best Local Similarity 20.0%; Pred. No. 3.9;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
CY 4 CXXXXXXXXX 13
DB 233 CTTATATATC 242
XX
RESULT 20
ADP31321
ID ADP31321 standard; protein; 270 AA.
XX
XX ADP31321;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2088.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
XX MO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003MO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SRP-2002; 2002US-0410946P.
PR 17-SRP-2002; 2002US-0410947P.
PR 17-SRP-2002; 2002US-0410948P.
PR 17-SRP-2002; 2002US-0410949P.
PR 17-SRP-2002; 2002US-0410953P.
PR 17-SRP-2002; 2002US-0410957P.
PR 17-SRP-2002; 2002US-0410958P.
PR 17-SRP-2002; 2002US-0410959P.
PR 17-SRP-2002; 2002US-0410960P.
PR 17-SRP-2002; 2002US-0410961P.
PR 17-SRP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 02-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493373P.
PR 08-AUG-2003; 2003US-0493577P.
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D;
XX Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;
XX Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI, 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3319; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cytostatic,
XX CC antitumorigenic, immunosuppressive, antibacterial and antiviral. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPOMB and is not in the specification.
XX
XX Sequence 270 AA;

Oy	4	CXXXXXXXXC 13
Db	245	CTTATTATAC 254
	RESULT 21	
	ADP31473	
ID	ADP31473	standard; protein; 294 AA.
XX		
AC	ADP31473;	
XX		
DT	12-AUG-2004	(first entry)
XX		
DE	Human secreted protein SEQ ID #2240.	
XX		
KW	Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide	
KM	cancer; inflammatory; immune; human secreted protein.	
XX		
OS	Homo sapiens.	
XX		
PN	WO2004035732-A2.	
XX		
PD	29-APR-2004.	
XX		
PF	28-AUG-2003;	2003WO-US026780.
XX		
PR	29-AUG-2002;	2002US-0406576P.
PR	29-AUG-2002;	2002US-0406579P.
PR	29-AUG-2002;	2002US-0406585P.
PR	29-AUG-2002;	2002US-0406588P.
PR	29-AUG-2002;	2002US-0406608P.
PR	29-AUG-2002;	2002US-0406611P.
PR	29-AUG-2002;	2002US-0406612P.
PR	29-AUG-2002;	2002US-0406616P.
PR	29-AUG-2002;	2002US-0406640P.
PR	29-AUG-2002;	2002US-0406642P.
PR	29-AUG-2002;	2002US-0406646P.
PR	29-AUG-2002;	2002US-0406653P.
PR	29-AUG-2002;	2002US-0406655P.
PR	29-AUG-2002;	2002US-0406666P.
PR	17-SEP-2002;	2002US-0410946P.
PR	17-SEP-2002;	2002US-0410947P.
PR	17-SEP-2002;	2002US-0410948P.
PR	17-SEP-2002;	2002US-0410949P.
PR	17-SEP-2002;	2002US-0410953P.
PR	17-SEP-2002;	2002US-0410958P.
PR	17-SEP-2002;	2002US-0410957P.
PR	17-SEP-2002;	2002US-0410958P.
PR	17-SEP-2002;	2002US-0410959P.
PR	17-SEP-2002;	2002US-0410960P.
PR	17-SEP-2002;	2002US-0410961P.
PR	17-SEP-2002;	2002US-0410962P.
PR	17-SEP-2002;	2002US-0411019P.
PR	17-SEP-2002;	2002US-0411022P.
PR	17-SEP-2002;	2002US-0411023P.
PR	17-SEP-2002;	2002US-0411024P.
PR	17-SEP-2002;	2002US-0411032P.
PR	17-SEP-2002;	2002US-0411035P.
PR	17-SEP-2002;	2002US-0411037P.
PR	17-SEP-2002;	2002US-0411041P.
PR	17-SEP-2002;	2002US-0411045P.
PR	17-SEP-2002;	2002US-0411046P.
PR	17-SEP-2002;	2002US-0411048P.
PR	17-SEP-2002;	2002US-0411052P.
PR	17-SEP-2002;	2002US-0411055P.
PR	17-SEP-2002;	2002US-0411073P.
PR	17-SEP-2002;	2002US-0411082P.
PR	17-SEP-2002;	2002US-0411101P.
PR	17-SEP-2002;	2002US-0411111P.
PR	18-APR-2003;	2003US-0463708P.
PR	18-APR-2003;	2003US-0463716P.
PR	18-APR-2003;	2003US-0463732P.
PR	02-MAY-2003;	2003US-0463719P.

PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471306P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 14-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 15-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PA Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 3471; 428bp; English.

PS The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.

XX Sequence 294 AA;

SO Query Match 56.2%; Score 18; DB 8; Length 294;

Best Local Similarity 20.0%; Pred. No. 4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 72 CATMAATC 81

RESULT 22

ID ADP31192 standard; protein; 297 AA.

XX ADP31192;

XX 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #1959.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

OS WO2004035732-A2.

XX

PD 29-APR-2004.
XX 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471306P.
PR 22-MAY-2003; 2003US-0472420P.
PR 09-JUN-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3190; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
SQ Sequence 297 AA;
Query Match 56.2%; Score 18; DB 8; Length 297;
Best Local Similarity 20.0%; Pred. No. 4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXC 13
Db 166 CTTATTATTC 175
RESULT 23
AAM25285
ID AAM25285 standard; protein; 328 AA.
XX
XX AAM25285;
AC
XX
XX 16-OCT-2001 (first entry)
DE Human protein sequence SEQ ID NO:800.
XX
XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiac; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnery; antitumor; osteopathic; eczema;
KW dermatological; antiallergic; antiaesthetic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antinaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder.
XX
XX Homo sapiens.
OS
XX
XX WO200153455-A2.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 22-DEC-2000; 2000WO-US035017.
PP
XX
XX 23-DEC-1999; 99US-00471275.
PR
XX 21-JAN-2000; 2000US-00488725.
PR
XX 25-APR-2000; 2000US-00552317.
PR
XX
XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;
PI WPI; 2001-457603/49.
DR N-PSDB; AAH99226.
XX
PT Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX
XX Claim 20; Page 187; 1217bp; English.
XX
XX AAH9166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and cells
CC they are expressed in, such as: antiinflammatory; antirheumatic;
CC antiallergic; immunosuppressive; antibacterial; endocrine; cardiac;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
CC antitumor; osteopathic; dermatological; antiallergic; antiaesthetic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders
SQ Sequence 328 AA;
Query Match 56.2%; Score 18; DB 4; Length 328;
Best Local Similarity 20.0%; Pred. No. 4.1;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXC 13
Db 63 CAAAAAAAC 72
RESULT 24
ADP30702
ID ADP30702 standard; protein; 339 AA.
XX
XX ADP30702;
AC
XX
XX 12-AUG-2004 (first entry)
DE Human secreted protein SEQ ID #1469.
XX
XX Cytostatic; Antiinflammatory; immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
OS
XX
XX WO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PP
XX
XX 29-AUG-2002; 2002US-0406576P.
PR
XX 29-AUG-2002; 2002US-0406579P.
PR
XX 29-AUG-2002; 2002US-0406585P.
PR
XX 29-AUG-2002; 2002US-0406588P.
PR
XX 29-AUG-2002; 2002US-0406608P.
PR
XX 29-AUG-2002; 2002US-0406611P.
PR
XX 29-AUG-2002; 2002US-0406612P.
PR
XX 29-AUG-2002; 2002US-0406616P.

PR 17-SEP-2002; 2002US-0411035P.
 PR 17-SEP-2002; 2002US-0411037P.
 PR 17-SEP-2002; 2002US-0411041P.
 PR 17-SEP-2002; 2002US-0411045P.
 PR 17-SEP-2002; 2002US-0411046P.
 PR 17-SEP-2002; 2002US-0411048P.
 PR 17-SEP-2002; 2002US-0411052P.
 PR 17-SEP-2002; 2002US-0411055P.
 PR 17-SEP-2002; 2002US-0411057P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411101P.
 PR 17-SEP-2002; 2002US-0411111P.
 PR 18-APR-2003; 2003US-0463700P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 02-MAY-2003; 2003US-0467206P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 22-MAY-2003; 2003US-0472430P.
 PR 09-JUN-2003; 2003US-0476609P.
 PR 09-JUN-2003; 2003US-0476641P.
 PR 08-JUL-2003; 2003US-0485218P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485224P.
 PR 08-JUL-2003; 2003US-0485225P.
 PR 14-JUL-2003; 2003US-0486446P.
 PR 14-JUL-2003; 2003US-0486480P.
 PR 15-JUL-2003; 2003US-0486891P.
 PR 15-JUL-2003; 2003US-0486960P.
 PR 08-AUG-2003; 2003US-0493341P.
 PR 08-AUG-2003; 2003US-0493370P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.
 XX
 XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
 XX
 XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
 PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
 XX
 XX WPI; 2004-348438/32.
 XX
 XX New nucleic acid molecule for diagnosing, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 PT genetic, bacterial and viral diseases.
 XX
 XX Claim 1; SEQ ID NO 3439; 428bp; English.
 XX
 XX The present invention relates to an isolated nucleic acid molecule
 CC encoding a polypeptide which is believed to be cytostatic,
 CC antiinflammatory, immunosuppressive, antibacterial and antiviral. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is
 CC available on WIPOMB and is not in the specification.
 XX
 XX Sequence 348 AA;
 SQ

Query Match 56.2%; Score 18; DB 8; Length 348;
 Best Local Similarity 20.0%; Pred. No. 4.1;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 25 CTTAATTAAC 34

RESULT 26
 ABG21039
 ID ABG21039 standard; protein; 356 AA.
 XX
 XX
 AC ABG21039;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 XX Novel human diagnostic protein #21030.
 DE
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 XX 30-MAR-2001; 2001WO-US008631.
 PF
 XX 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Drmanac RT, Liu C, Tang YT;
 PI
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS85226.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 XX Claim 20; SEQ ID NO 51398; 103bp; English.
 PS
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 356 AA;
 SQ

Query Match 56.2%; Score 18; DB 4; Length 356;
 Best Local Similarity 20.0%; Pred. No. 4.2;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 183 CAAAAAASC 192

RESULT 27
 ADP31267

ID ADP31267 standard; protein; 357 AA.
 AC ADP31267;
 XX
 XX
 DT 12-AUG-2004 (first entry)
 DE Human secreted protein SEQ ID #2034.
 DE
 XX
 KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
 KW cancer; inflammatory; immune; human secreted protein.
 OS Homo sapiens.
 XX
 XX
 PN MO2004035732-A2.
 XX
 PD 29-APR-2004.
 XX
 XX
 PF 28-AUG-2003; 2003WO-US026780.
 XX
 XX
 PR 29-AUG-2002; 2002US-0406576P.
 PR 29-AUG-2002; 2002US-0406579P.
 PR 29-AUG-2002; 2002US-0406585P.
 PR 29-AUG-2002; 2002US-0406588P.
 PR 29-AUG-2002; 2002US-0406608P.
 PR 29-AUG-2002; 2002US-0406611P.
 PR 29-AUG-2002; 2002US-0406612P.
 PR 29-AUG-2002; 2002US-0406616P.
 PR 29-AUG-2002; 2002US-0406640P.
 PR 29-AUG-2002; 2002US-0406642P.
 PR 29-AUG-2002; 2002US-0406646P.
 PR 29-AUG-2002; 2002US-0406653P.
 PR 29-AUG-2002; 2002US-0406655P.
 PR 29-AUG-2002; 2002US-0406666P.
 PR 17-SEP-2002; 2002US-0410946P.
 PR 17-SEP-2002; 2002US-0410947P.
 PR 17-SEP-2002; 2002US-0410948P.
 PR 17-SEP-2002; 2002US-0410949P.
 PR 17-SEP-2002; 2002US-0410953P.
 PR 17-SEP-2002; 2002US-0410957P.
 PR 17-SEP-2002; 2002US-0410958P.
 PR 17-SEP-2002; 2002US-0410959P.
 PR 17-SEP-2002; 2002US-0410960P.
 PR 17-SEP-2002; 2002US-0410961P.
 PR 17-SEP-2002; 2002US-0410962P.
 PR 17-SEP-2002; 2002US-0411019P.
 PR 17-SEP-2002; 2002US-0411022P.
 PR 17-SEP-2002; 2002US-0411023P.
 PR 17-SEP-2002; 2002US-0411024P.
 PR 17-SEP-2002; 2002US-0411032P.
 PR 17-SEP-2002; 2002US-0411035P.
 PR 17-SEP-2002; 2002US-0411037P.
 PR 17-SEP-2002; 2002US-0411041P.
 PR 17-SEP-2002; 2002US-0411045P.
 PR 17-SEP-2002; 2002US-0411046P.
 PR 17-SEP-2002; 2002US-0411048P.
 PR 17-SEP-2002; 2002US-0411052P.
 PR 17-SEP-2002; 2002US-0411055P.
 PR 17-SEP-2002; 2002US-0411073P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411101P.
 PR 17-SEP-2002; 2002US-0411111P.
 PR 18-APR-2003; 2003US-0463700P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 02-MAY-2003; 2003US-0467230P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 19-MAY-2003; 2003US-0471366P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.
 PR 09-JUN-2003; 2003US-0476641P.
 PR 08-JUL-2003; 2003US-0485218P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485224P.
 PR 08-JUL-2003; 2003US-0485225P.
 PR 14-JUL-2003; 2003US-0486446P.
 PR 14-JUL-2003; 2003US-0486480P.
 PR 15-JUL-2003; 2003US-0486891P.
 PR 15-JUL-2003; 2003US-0486960P.
 PR 08-AUG-2003; 2003US-0493341P.
 PR 08-AUG-2003; 2003US-0493370P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.
 XX
 XX
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
 XX
 XX
 PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
 PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
 XX
 DR WPI; 2004-348438/32.
 XX
 XX
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 PT genetic, bacterial and viral diseases.
 XX
 PS Claim 1; SEQ ID NO 3265; 428bp; English.
 XX
 XX
 CC The present invention relates to an isolated nucleic acid molecule
 CC encoding a polypeptide which is believed to be cytosolic,
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is
 CC available on WIPOMEB and is not in the specification.
 XX
 SO Sequence 357 AA;
 QY 4 CXXXXXXXXC 13
 Db 14 CATATATATAC 23
 RESULT 28
 ID ADP30505 standard; protein; 357 AA.
 XX
 AC ADP30505;
 XX
 DT 12-AUG-2004 (first entry)
 DE Human secreted protein SEQ ID #1272.
 DE
 XX
 KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
 KW cancer; inflammatory; immune; human secreted protein.
 OS Homo sapiens.
 XX
 XX
 PN MO2004035732-A2.
 PD 29-APR-2004.
 XX
 PF 28-AUG-2003; 2003WO-US026780.
 XX
 XX
 PR 29-AUG-2002; 2002US-0406576P.
 PR 29-AUG-2002; 2002US-0406579P.
 PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471366P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.

XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 2503; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC anti-inflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 357 AA;
XX
Query Match 56.2%; Score 18; DB 8; Length 357;
Best Local Similarity 20.0%; Pred. No. 4.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXC 13
DB 169 CTAATTTC 178
RESULT 29
ADP31439
ID ADP31439 standard; protein; 360 AA.
XX
AC ADP31439;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2206.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX PT genetic, bacterial and viral diseases.
XX
XX
XX Claim 1; SEQ ID NO 3437; 428bp; English.
XX
XX
XX The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cytostatic,
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPOMB and is not in the specification.
XX
XX
XX Sequence 360 AA:
Query Match 56.2%; Score 18; DB 8; Length 360;
Best Local Similarity 20.0%; Pred. No. 4.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
Db 70 CTAATTAAC 79
RESULT 30
ADP31218
ADP31218 standard; protein; 390 AA.
XX
XX ADP31218;
XX
XX 12-AUG-2004 (first entry)
XX
XX
XX Human secreted protein SEQ ID #1985.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; Inflammatory; immune; human secreted protein.
OS
XX Homo sapiens.
XX
XX MO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411101P.
XX 17-SEP-2002; 2002US-0411111P.
XX 18-APR-2003; 2003US-0463700P.
XX 18-APR-2003; 2003US-0463708P.
XX 18-APR-2003; 2003US-0463716P.
XX 18-APR-2003; 2003US-0463732P.
XX 02-MAY-2003; 2003US-0467199P.

```

PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Hsiehan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI, 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PS genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3216; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic.
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 390 AA;
XX
Query Match 56.2%; Score 18; DB 8; Length 390;
Best Local Similarity 20.0%; Pred. No. 4,2; Mismatches 0; Gaps 0.
Matches 2; Conservative 0; Indels 0;
OY 4 CXXXXXXXXXC 13
DB |
60 CAAAAAATTC 69
RESULT 31
ADP31159
ID ADP31159 standard; protein; 421 AA.
AC ADP31159;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1926.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
```

FD	29-APR-2004.	
XX		
PR	28-AUG-2003,	2003MO-US026780.
XX		
PR	29-AUG-2002,	2002US-0406576P.
PR	29-AUG-2002,	2002US-0406578P.
PR	29-AUG-2002,	2002US-0406585P.
PR	29-AUG-2002,	2002US-0406588P.
PR	29-AUG-2002,	2002US-0406608P.
PR	29-AUG-2002,	2002US-0406611P.
PR	29-AUG-2002,	2002US-0406612P.
PR	29-AUG-2002,	2002US-0406616P.
PR	29-AUG-2002,	2002US-0406642P.
PR	29-AUG-2002,	2002US-0406644P.
PR	29-AUG-2002,	2002US-0406653P.
PR	29-AUG-2002,	2002US-0406655P.
PR	29-AUG-2002,	2002US-0406665P.
PR	29-AUG-2002,	2002US-0406666P.
PR	17-SEP-2002,	2002US-04109347P.
PR	17-SEP-2002,	2002US-0410947P.
PR	17-SEP-2002,	2002US-0410947P.
PR	17-SEP-2002,	2002US-0410960P.
PR	17-SEP-2002,	2002US-0410961P.
PR	17-SEP-2002,	2002US-0410962P.
PR	17-SEP-2002,	2002US-0411019P.
PR	17-SEP-2002,	2002US-0411022P.
PR	17-SEP-2002,	2002US-0411024P.
PR	17-SEP-2002,	2002US-0411024P.
PR	17-SEP-2002,	2002US-0411032P.
PR	17-SEP-2002,	2002US-0411035P.
PR	17-SEP-2002,	2002US-0411037P.
PR	17-SEP-2002,	2002US-0411041P.
PR	17-SEP-2002,	2002US-0411045P.
PR	17-SEP-2002,	2002US-0411046P.
PR	17-SEP-2002,	2002US-0411104P.
PR	17-SEP-2002,	2002US-0411108P.
PR	17-SEP-2002,	2002US-0411052P.
PR	17-SEP-2002,	2002US-0411055P.
PR	17-SEP-2002,	2002US-0411073P.
PR	17-SEP-2002,	2002US-0411082P.
PR	17-SEP-2002,	2002US-0411101P.
PR	17-SEP-2002,	2002US-0411111P.
PR	18-APR-2003,	2003US-04063700P.
PR	18-APR-2003,	2003US-04063708P.
PR	18-APR-2003,	2003US-04063718P.
PR	18-APR-2003,	2003US-04063732P.
PR	18-APR-2003,	2003US-04063739P.
PR	02-MAY-2003,	2003US-04067199P.
PR	02-MAY-2003,	2003US-04067201P.
PR	02-MAY-2003,	2003US-04067203P.
PR	02-MAY-2003,	2003US-04067230P.
PR	19-MAY-2003,	2003US-0407136P.
PR	19-MAY-2003,	2003US-0407136P.
PR	22-MAY-2003,	2003US-04072430P.
PR	09-JUN-2003,	2003US-04076509P.
PR	09-JUN-2003,	2003US-04076614P.
PR	08-JUL-2003,	2003US-04085218P.
PR	08-JUL-2003,	2003US-04085223P.
PR	08-JUL-2003,	2003US-04085223P.
PR	08-JUL-2003,	2003US-04085446P.
PR	14-JUL-2003,	2003US-04086480P.
PR	14-JUL-2003,	2003US-04086486P.
PR	15-JUL-2003,	2003US-04086961P.
PR	15-JUL-2003,	2003US-04086961P.
PR	08-AUG-2003,	2003US-04093414P.
PR	08-AUG-2003,	2003US-04093370P.
PR	08-AUG-2003,	2003US-04093573P.
XX	08-AUG-2003,	2003US-04093577P.

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX MPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
PS
PS Claim 1; SEQ ID NO 3157; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 421 AA;
XX
Query Match 56.2%; Score 18; DB 8; Length 421;
Best Local Similarity 20.0%; Pred. No. 4.3;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXC 13
Db 385 CAAAAAATC 394
XX
RESULT 32
ADP31495
ID ADP31495 standard; protein; 426 AA.
XX
XX ADP31495;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2262.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune, human secreted protein.
XX
OS Homo sapiens.
XX
XX
XX MO2004035732-A2.
XX
XX
XX PD 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX
XX 29-AUG-2002; 2002US-0406579P.
XX
XX 29-AUG-2002; 2002US-0406585P.
XX
XX 29-AUG-2002; 2002US-0406588P.
XX
XX 29-AUG-2002; 2002US-0406608P.
XX
XX 29-AUG-2002; 2002US-0406611P.
XX
XX 29-AUG-2002; 2002US-0406612P.
XX
XX 29-AUG-2002; 2002US-0406616P.
XX
XX 29-AUG-2002; 2002US-0406640P.
XX
XX 29-AUG-2002; 2002US-0406642P.
XX
XX 29-AUG-2002; 2002US-0406646P.
XX
XX 29-AUG-2002; 2002US-0406653P.
XX
XX 29-AUG-2002; 2002US-0406655P.
XX
XX 29-AUG-2002; 2002US-0406666P.
XX
XX 17-SEP-2002; 2002US-0410946P.
XX
XX 17-SEP-2002; 2002US-0410947P.
XX
XX 17-SEP-2002; 2002US-0410948P.
XX
XX 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485244P.
PR 08-JUL-2003; 2003US-0485255P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486860P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
XX PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
XX PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX MPI; 2004-348438/32.
XX
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX
XX PS Claim 1; SEQ ID NO 3493; 428bp; English.
XX
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPOMB and is not in the specification.

```
XX Sequence 426 AA;
SQ
Query Match          56.2%; Score 18; DB 8; Length 426;
Best Local Similarity 20.0%; Pred. No. 4.3;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 66 CATTAAATTC 75

RESULT 33
ABO82561
ID ABO82561 standard; protein; 438 AA.
XX
AC ABO82561;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #14736.
XX
KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
XX
DR N-PSDB; ABD16132.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 31307; 455bp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 438 AA;
Query Match          56.2%; Score 18; DB 7; Length 438;
Best Local Similarity 20.0%; Pred. No. 4.3;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 66 CATTAAATTC 75

RESULT 34
ADP31465
ID ADP31465 standard; protein; 453 AA.
XX
AC ADP31465;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2232.
XX
KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 29-AUG-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411028P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411033P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
```


PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 03-JUN-2003; 2003US-0476609P.
PR 03-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,
XX
DR MPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 3463; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 453 AA;
Query Match 56.2%; Score 18; DB 8; Length 453;
Best Local Similarity 20.0%; Pred. No. 4.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXX 13
Db 312 CAAATTATTC 321
RESULT 35
ADP30854
ID ADP30854 standard; protein; 471 AA.
XX
AC ADP30854;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1621.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX

PF 28-AUG-2003; 2003MO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX

Query Match 56.2%; Score 18; DB 8; Length 525;
Best Local Similarity 20.0%; Pred. No. 4.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
Db 475 CAAATTTAC 484

RESULT 37
ADP30855
ID ADP30855 standard; protein; 549 AA.
XX
AC ADP30855;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1622.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN MO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485244P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beauregard PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 2853; 428bp; English.
XX
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
XX
SQ Sequence 549 AA;

Query Match 56.2%; Score 18; DB 8; Length 549;
Best Local Similarity 20.0%; Pred. No. 4.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
Db 220 CAAATTTATC 229

RESULT 38
ADP31416
ID ADP31416 standard; protein; 555 AA.
XX
AC ADP31416;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2183.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.

	XX	29-AUG-2002;	2002US-0406666P.
PR	17-SEP-2002;	2002US-0410946P.	
PR	17-SEP-2002;	2002US-0410947P.	
PR	17-SEP-2002;	2002US-0410948P.	
PR	17-SEP-2002;	2002US-0410949P.	
PR	17-SEP-2002;	2002US-0410953P.	
PR	17-SEP-2002;	2002US-0410957P.	
PR	17-SEP-2002;	2002US-0410958P.	
PR	17-SEP-2002;	2002US-0410959P.	
PR	17-SEP-2002;	2002US-0410960P.	
PR	17-SEP-2002;	2002US-0410961P.	
PR	17-SEP-2002;	2002US-0410962P.	
PR	17-SEP-2002;	2002US-0411019P.	
PR	17-SEP-2002;	2002US-0411022P.	
PR	17-SEP-2002;	2002US-0411023P.	
PR	17-SEP-2002;	2002US-0411024P.	
PR	17-SEP-2002;	2002US-0411032P.	
PR	17-SEP-2002;	2002US-0411035P.	
PR	17-SEP-2002;	2002US-0411037P.	
PR	17-SEP-2002;	2002US-0411041P.	
PR	17-SEP-2002;	2002US-0411045P.	
PR	17-SEP-2002;	2002US-0411046P.	
PR	17-SEP-2002;	2002US-0411052P.	
PR	17-SEP-2002;	2002US-0411054P.	
PR	17-SEP-2002;	2002US-0411055P.	
PR	17-SEP-2002;	2002US-0411073P.	
PR	17-SEP-2002;	2002US-0411082P.	
PR	17-SEP-2002;	2002US-0411101P.	
PR	17-SEP-2002;	2002US-0411111P.	
PR	18-APR-2003;	2003US-0463700P.	
PR	18-APR-2003;	2003US-0463708P.	
PR	18-APR-2003;	2003US-0463716P.	
PR	18-APR-2003;	2003US-0463732P.	
PR	02-MAY-2003;	2003US-0467199P.	
PR	02-MAY-2003;	2003US-0467201P.	
PR	02-MAY-2003;	2003US-0467203P.	
PR	02-MAY-2003;	2003US-0467203P.	
PR	19-MAY-2003;	2003US-0471306P.	
PR	19-MAY-2003;	2003US-0471336P.	
PR	22-MAY-2003;	2003US-0472420P.	
PR	22-MAY-2003;	2003US-0472430P.	
PR	09-JUN-2003;	2003US-0476609P.	
PR	09-JUN-2003;	2003US-0476641P.	
PR	08-JUL-2003;	2003US-0485218P.	
PR	08-JUL-2003;	2003US-0485223P.	
PR	08-JUL-2003;	2003US-0485224P.	
PR	08-JUL-2003;	2003US-0485325P.	
PR	14-JUL-2003;	2003US-0486446P.	
PR	14-JUL-2003;	2003US-0486480P.	
PR	15-JUL-2003;	2003US-0486891P.	
PR	15-JUL-2003;	2003US-0486960P.	
PR	08-AUG-2003;	2003US-0493341P.	
PR	08-AUG-2003;	2003US-0493370P.	
PR	08-AUG-2003;	2003US-0493573P.	
PR	08-AUG-2003;	2003US-0493577P.	
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.		
XX			
PI	Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,		
PI	Halenbeck RP, Huang WM, Kotthakota S, Haislan L, Linnemann T,		
PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;		
XX	WPI; 2004-346438/32.		
DR			
XX			
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases		
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,		
PT	genetic, bacterial and viral diseases.		
XX			
PS	Claim 1; SEQ ID NO 3415; 428bp; English.		
CC	The present invention relates to an isolated nucleic acid molecule		
CC	encoding a polypeptide which is believed to be cytostatic,		
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The		

composition and methods are useful for diagnosing, preventing and treating diseases such as proliferative (e.g. cancer), inflammatory, immune, metabolic, genetic, bacterial and viral diseases. The present sequence represents a human secreted protein. The present sequence is available on WIPOMED and is not in the specification.

Sequence 555 AA;

Query Match	56.2%	Score 18;	DB 8;	Length 555;
Best Local Similarity	20.0%;	Pred. No. 4.6;		
Matches	2;	Conservative	0;	Mismatches 8;
			Indels	0;
			Gaps	0;

QY	4	XXXXXXXXXX	13
Db	211	CTTTATTAC	220

RESULT 40
ADP31699
ID ADP31699 standard; protein; 588 AA

DT 12-AUG-2004 (first entry)

DE	Human secreted protein SEQ ID #2466
DE	Human secreted protein SEQ ID #2466

KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide
 KN cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

PN WO2004035732-A2.

PD 29-APR-2004.

PF 28-AUG-2003; 2003WO-US026780.

PR 29-AUG-2002; 2002US-0406576P.

PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-040664OP.

PR 29-AUG-2002; 2002US-0406646P.

PR 29-AUG-2002; 2002US-0406655P.

PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410953P.

PR 17-SEP-2002; 2002US-0410958P.

PR 17-SEP-2002; 2002US-0410960P.

PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411022P.

PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411023P.

PR 17-SEP-2002; 2002US-0411035P.
PR 17 SEP 2002 2002TTC 041103ZD

PR 17-SEP-2002; 2002US-0411041P.
PP 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.

CC represents a non TM region of a protein encoded by a cDNA splice variant
CC from the KIAA0779 gene.
XX
SQ Sequence 591 AA:
Query Match 56.2%; Score 18; DB 9; Length 591;
Best Local Similarity 20.0%; Pred. No. 4.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXX 13
DB 173 CAAAAAAAAAC 182
RESULT 42
ADZ56223
ID ADZ56223 standard; protein; 591 AA.
XX
AC ADZ56223;
XX
DT 30-JUN-2005 (first entry)
XX
DE Human KIAA0779 splice variant clone 24980850:24980849a, nonTM_1 #2.
XX
KM KIAA0779; gene expression; inflammation; antiinflammatory; cancer;
KM proliferation; neoplasm; cytostatic; immune disorder; immunomodulator;
KM metabolic disorder; metabolic; viral infection; virucide; infection.
XX
OS Homo sapiens.
XX
PN MO2005035569-A2.
XX
PD 21-APR-2005.
XX
PF 12-OCT-2004; 2004MO-US033408.
XX
PR 10-OCT-2003; 2003US-0510612P.
XX
PA (PIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Wong JGP, Hestir K, Collins ALT;
XX
DR WPI; 2005-296268/30.
XX
PT New isolated KIAA0779 nucleic acids and polypeptides, useful for
PT diagnosing, preventing and/or treating inflammatory, immune, viral
PT disorders and cancer, such as kidney, lung, ovarian, bladder, breast,
PT prostate and skin cancers.
XX
PS Claim 11, SEQ ID NO 28; 121bp; English.
XX
CC The invention relates to an isolated nucleic acid molecule comprises at
CC least one polynucleotide sequence (appearing as ADZ56196-ADZ56197,
CC ADZ56199-ADZ56200, ADZ56226-ADZ56227 and ADZ56230-ADZ56231), sequences
CC hybridizing to them under high stringency conditions, sequences having at
CC least 80% sequence identity to them, their complements or biologically
CC active fragments. The nucleic acids are splice variants of the human
CC KIAA0779 gene, encoding transmembrane domain protein(s). Also included
CC are a double-stranded isolated nucleic acid molecule comprising the
CC nucleic acid molecule cited above, a vector comprising the isolated
CC nucleic acid molecule cited above, (and a promoter that regulates the
CC expression of the nucleic acid molecule), a recombinant host cell
CC comprising the nucleic acid molecule cited above, an isolated polypeptide
CC (comprising the encoded proteins or the non-transmembrane (TM) regions of
CC the proteins), a method of making a recombinant host cell, a method of
CC making a polypeptide, a method of determining the presence of the nucleic
CC acid molecule cited above in a sample, a method of determining the
CC presence of a specific antibody to the polypeptide of (4) in a sample, a
CC method of determining the presence of the polypeptide in a sample, an
CC antibody specifically binding to and/or interfering with the biological
CC activity of the nucleic acid molecule cited above (or the polypeptide or
CC its biologically active fragment), a composition comprising a
CC pharmaceutical carrier or excipient (and one or more active agents chosen

CC from the nucleic acid molecule cited above, the vector, the polypeptide,
CC and the antibody), a bacteriophage comprising the antibody (or its
CC fragment), a bacterial cell comprising the bacteriophage, a recombinant
CC host cell that produces the antibody, an animal injected with one or more
CC active agents (chosen from the nucleic acid molecule, the vector, the
CC host cell, the polypeptide, and the antibody), a diagnostic kit
CC (comprising a nucleic acid molecule having at least 6 contiguous
CC nucleotides from the nucleic acid molecule cited above, the isolated
CC polypeptide, or antibody, and reagents to carry out an immunoassay), a
CC method of making an antibody, a method of identifying a modulating agent
CC that modulates the biological activity of the polypeptide, a modulator
CC composition comprising a modulator and a pharmaceutical carrier, a method
CC of treating a disease in a subject, and a method of treating cancer,
CC preferably kidney, cervical, squamous lung, ovarian, bladder, breast,
CC endometrial, prostate, and skin cancer, in a subject. The methods and
CC compositions of the present invention are useful for the diagnosis,
CC prevention and/or treatment of inflammatory, immune, viral disorders and
CC cancer, such as kidney, cervical, squamous lung, ovarian, bladder,
CC breast, endometrial, prostate, and skin cancer. The present sequence
CC represents a non TM region of a protein encoded by a cDNA splice variant
CC from the KIAA0779 gene.
XX
SQ Sequence 591 AA:
Query Match 56.2%; Score 18; DB 9; Length 591;
Best Local Similarity 20.0%; Pred. No. 4.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXX 13
DB 173 CAAAAAAAAAC 182
RESULT 43
ADP30865
ID ADP30865 standard; protein; 600 AA.
XX
AC ADP30865;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1632.
XX
DE Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN MO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003MO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.

PR 17-SEP-2002; 2002US-041111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halembeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.

DR New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.

PS Claim 1; SEQ ID NO 2938; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.

SO Sequence 604 AA;

Query Match 56.2%; Score 18; DB 8; Length 604;

Best Local Similarity 20.0%; Pred. No. 4.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 430 CTTATTAAAC 439

RESULT 45

ID ADP30941 standard; protein; 604 AA.

XX ADP30941;

DT 12-AUG-2004 (first entry)

XX Human secreted protein SEQ ID #1708.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

XX MO2004035732-A2.

PN 29-APR-2004.

PD 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 2939; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC anti-inflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
SQ Sequence 604 AA:
Query Match 56.2%; Score 18; DB 8; Length 604;
Best Local Similarity 20.0%; Pred. No. 4.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXX 13
DB 430 CTTATTAAAC 439
RESULT 46
ADP30507
ID ADP30507 standard; protein; 605 AA.
XX
XX ADP30507;
AC
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1274.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX
XX
PD 29-APR-2004.
XX
XX 26-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-048518P.
PR 08-JUL-2003; 2003US-0485233P.
PR 08-JUL-2003; 2003US-048524P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2505; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,

CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.

XX
SQ Sequence 605 AA;

Query Match 56.2%; Score 18; DB 8; Length 605;
Best Local Similarity 20.0%; Pred. No. 4.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 570 CAAATTATTC 579

RESULT 47

ID ADP31657 standard; protein; 617 AA.

XX ADP31657;

XX 12-AUG-2004 (first entry)

XX Human secreted protein SEQ ID #2424.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

XX MO2004035732-A2.

XX 29-APR-2004.

XX 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

XX 29-AUG-2002; 2002US-0406666P.

XX 17-SEP-2002; 2002US-0410946P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410958P.

XX 17-SEP-2002; 2002US-0410959P.

XX 17-SEP-2002; 2002US-0410960P.

XX 17-SEP-2002; 2002US-0410961P.

XX 17-SEP-2002; 2002US-0410962P.

XX 17-SEP-2002; 2002US-0411019P.

XX 17-SEP-2002; 2002US-0411022P.

XX 17-SEP-2002; 2002US-0411023P.

XX 17-SEP-2002; 2002US-0411024P.

XX 17-SEP-2002; 2002US-0411032P.

XX 17-SEP-2002; 2002US-0411035P.

XX 17-SEP-2002; 2002US-0411037P.

XX 17-SEP-2002; 2002US-0411041P.

XX 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.

PR 17-SEP-2002; 2002US-0411048P.

PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.

PR 17-SEP-2002; 2002US-0411101P.

PR 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463716P.

PR 18-APR-2003; 2003US-0463732P.

PR 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.

PR 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.

PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.

PR 09-JUN-2003; 2003US-0476641P.

PR 08-JUL-2003; 2003US-0485218P.

PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485224P.

PR 14-JUL-2003; 2003US-0486452P.

PR 14-JUL-2003; 2003US-0486466P.

PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486891P.

PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.

PR 08-AUG-2003; 2003US-0493577P.

XX (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LR, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;

PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Lunemann T;

PI Pierce K, Wang Y, Wong JCF, Wu G, Zhang H;

DR WPI; 2004-348438/32.

PT New nucleic acid molecule for diagnosing, preventing or treating diseases

PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,

PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 3655; 428pp; English.

XX The present invention relates to an isolated nucleic acid molecule

XX encoding a polypeptide which is believed to be cytostatic.

XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The

XX composition and methods are useful for diagnosing, preventing and

XX treating diseases such as proliferative (e.g. cancer), inflammatory,

XX immune, metabolic, genetic, bacterial and viral diseases. The present

XX sequence represents a human secreted protein. The present sequence is

XX available on WIPOMEB and is not in the specification.

XX Sequence 617 AA;

Query Match 56.2%; Score 18; DB 8; Length 617;
Best Local Similarity 20.0%; Pred. No. 4.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 67 CAAATTATAC 76

RESULT 48

ID ABR41596 standard; protein; 626 AA.

XX ABR41596

AC ABR41596;
XX
DT 02-JUN-2003 (first entry)
XX
DE Human DITHP antigen recognition protein.
XX
KM Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
KM cancer; cell proliferative disorder; autoimmune disorder;
KM inflammatory disorder; infection; hormonal disorder; metabolic disorder;
KM neurological disorder; gastrointestinal disorder; transport disorder;
KM connective tissue disorder; drug screening; proteome analysis;
KM gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
KM disease model; toxicological testing; transcript imaging;
KM antigen recognition.
XX
OS Homo sapiens.
XX
PN WO200297031-A2.
XX
PD 05-DEC-2002.
XX
PF 27-MAR-2002; 2002WO-US010056.
XX
PR 28-MAR-2001; 2001US-0279619P.
PR 29-MAR-2001; 2001US-0280067P.
PR 29-MAR-2001; 2001US-0280068P.
PR 16-MAY-2001; 2001US-0291820P.
PR 17-MAY-2001; 2001US-0291829P.
PR 17-MAY-2001; 2001US-0291849P.
PR 19-JUN-2001; 2001US-0299428P.
PR 20-JUN-2001; 2001US-0299776P.
PR 20-JUN-2001; 2001US-0300001P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Thason O, Yap PE, Amshay SR;
PI Daugherty SC, Dam TC, Nguyen DA, Kleefeld Y, Gerstin EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Ian RV, Urashka ME;
XX
DR MPI; 2003-129518/12.
DR N-PSDB; ACC646534.
XX
PT Novel human diagnostic and therapeutic polypeptide useful for identifying
PT test compound which specifically binds to a polypeptide encoded by human
PT diagnostic and therapeutic polynucleotide, and to induce antibodies.
XX
PS Claim 27; SEQ ID NO 1131; 591pp; English.
XX
XX The invention relates to novel human diagnostic and therapeutic
CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to
CC polynucleotide sequences at least 90% identical to the dithp cDNA
CC sequences of the invention; recombinant vectors, host cells and
CC transgenic organisms comprising a dithp nucleic acid sequence; the
CC recombinant production of DITHP proteins; antibodies specific for DITHP
CC proteins; microarrays comprising dithp nucleic acid sequences; methods of
CC detecting dithp nucleotide and protein sequences; methods of screening
CC for compounds which specifically bind a DITHP protein; and methods of
CC assessing the toxicity of test compounds using a dithp hybridisation
CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
CC diagnosis of a wide variety of conditions including cancer and other cell
CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
CC viral, fungal or parasitic infections; hormonal disorders; metabolic
CC disorders; neurological disorders; gastrointestinal disorders; transport
CC disorders; and connective tissue disorders. They may also be used to
CC screen for modulators of protein activity or gene expression. DITHP
CC proteins can additionally be used in analysis of the proteome of a tissue
CC or cell type and to induce antibodies. The dithp nucleic acids are
CC additionally useful in somatic or germline gene therapy of the disorders
CC mentioned above, as a source of antisense sequences, as a source of
CC probes and primers, in genotyping and identification of individuals, in

CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
CC The present sequence represents a DITHP protein which has antigen
CC recognition activity. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 626 AA;
XX
Query Match 56.2%; Score 18; DB 6; Length 626;
Best Local Similarity 20.0%; Pred. No. 4.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXC 13
DB 41 CTSTSSASC 50
XX
RESULT 49
ID ADP30858
AC ADP30858 standard; protein; 645 AA.
XX
AC ADP30858;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1625.
XX
KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.

17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471316P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476611P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RT, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

DR MPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
PT

PS Claim 1; SEQ ID NO 2856; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX

XX Sequence 645 AA;

Query Match 56.2%; Score 18; DB 8; Length 645;
Best Local Similarity 20.0%; Pred. NO. 4.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXX 13
DB 161 CAAATTTTTC 170

RESULT 50
AA94907
ID AAY94907 standard; protein; 653 AA.
XX

AC AAY94907;
XX
XX 16-JUN-2000 (first entry)
DT

DE Human secreted protein clone ca106_19x protein sequence SEQ ID NO:20.

XX Human; secreted protein; immunestimulant; immunosuppressant; virucide;
KW antibacterial; antifungal; cytostatic; antiinflammatory; dermatological;
KW antidiabetic; antiasthmatic; antiarthritic; antirheumatic; protozoacide;
KW antithyroid; immune deficiency; severe combined immunodeficiency; SCID;
KW infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;
KW connective tissue disease; multiple sclerosis; erythematosis;
KW rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;
KW Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;
KW insulin dependent diabetes mellitus; graft-versus-host-disease;
KW autoimmune inflammatory eye disease; allergy.

OS Homo sapiens.

PN MO200009552-A1.

PD 24-FEB-2000.

PF 13-AUG-1999; 99MO-US018298.

XX 14-AUG-1998; 98US-0096622P.
PR 17-AUG-1998; 98US-0096815P.
PR 04-SEP-1998; 98US-0099229P.
PR 23-OCT-1998; 98US-0105368P.
PR 08-JAN-1999; 99US-0115334P.
PR 12-FEB-1999; 99US-0119931P.
PR 18-FEB-1999; 99US-0120575P.
PR 30-APR-1999; 99US-0132020P.
PR 11-AUG-1999; 99US-0148424P.

(GEMV) GENETICS INST INC.

PA Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Agostino MJ, Steinger RJ, Spaulding V;
PI Wong GG, Clark HF, Reichel K;

DR MPI; 2000-205979/18.

XX New polynucleotides encoding secreted proteins, which may have e.g.
PT nutritional, chemokine, immune stimulating or suppressing, hematopoiesis
PT regulating, tissue growth, activin/inhibin antiinflammatory or tumor
PT inhibition activity.
XX

PS Claim 29; Page 487-489; 641pp; English.

XX AAI1618 to AAI1697 encode the human secreted proteins given in AAY94898
CC to AAY94980, isolated from human adult brain, adult thyroid, adult
CC retina, foetal carcinoma, adult blood, adult neural, foetal kidney, adult
CC placenta, adult testis, whole embryo, adult cartilage, kidney, foetal
CC brain, adult thymus, foetal placenta, adult uterus, adult tumour, and
CC adult bladder, cDNA libraries. The polynucleotides and proteins are
CC predicted to have biological activities which would make them suitable
CC for treating, preventing or ameliorating medical conditions in humans and
CC animals. The polynucleotides can be used as markers for tissues in which
CC the protein is preferentially expressed, as molecular weight markers on
CC Southern gels, and as chromosome markers or tags to identify chromosomes
CC or to map gene positions. The proteins can be used in the treatment of
CC immune deficiencies and disorders, such as severe combined
CC immunodeficiency (SCID), as well as viral, bacterial, fungal and other
CC infections. These infections include human immunodeficiency virus (HIV),
CC hepatitis, herpesviruses, mycobacteria, leishmania spp., malaria and
CC candidiasis. The proteins can be used to treat autoimmune disorders such
CC as connective tissue disease, multiple sclerosis, systemic lupus
CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation,
CC Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent
CC diabetes mellitus, myasthenia gravis, graft-versus-host-disease and
CC autoimmune inflammatory eye disease. The proteins can also be used to
CC treat allergic conditions, such as asthma. AAI1698 to AAI16774 represent

CC probes for the human secreted proteins from the present invention
XX
SQ Sequence 653 AA;

Query Match 56.2%; Score 18; DB 3; Length 653;
Best Local Similarity 20.0%; Pred. No. 4.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 13
|
Db 173 CAAAAAAC 182

Search completed: January 4, 2006, 15:56:03
Job time : 122.496 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 ; Search time 16.5565 Seconds
(without alignments)
92.983 Million cell updates/sec

Title: US-09-932-322-4
Perfect score: 32
Sequence: 1 XXXXXXXXXXXXXXXX 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	56.2	313	2	SS9448
2	18	56.2	480	2	E75433
3	17	53.1	194	2	JQ1560
4	17	53.1	220	2	JC4082
5	17	53.1	376	2	AB8290
6	17	53.1	388	2	C97694
7	17	53.1	389	2	T23167
8	17	53.1	570	2	T37314
9	17	53.1	634	2	S33575
10	17	53.1	942	2	D87803
11	17	53.1	2704	2	S09118
12	16	50.0	71	2	F84312
13	16	50.0	85	2	E70531
14	16	50.0	104	2	S51479
15	16	50.0	105	2	A71249
16	16	50.0	108	2	F72549
17	16	50.0	126	2	A23473
18	16	50.0	129	2	T49498
19	16	50.0	133	2	A96746
20	16	50.0	147	2	S09762
21	16	50.0	153	2	PN0103
22	16	50.0	161	2	T28088
23	16	50.0	164	2	JQ1252
24	16	50.0	229	2	T34277
25	16	50.0	242	2	A45724
26	16	50.0	244	2	S72219
27	16	50.0	251	2	A55035
28	16	50.0	262	1	JQ1724
29	16	50.0	263	2	S47537

30	16	50.0	275	2	A85856	probable elongatio
31	16	50.0	275	2	B64986	hypothetical 30.9
32	16	50.0	275	2	G91011	probable elongatio
33	16	50.0	294	2	T23682	hypothetical prote
34	16	50.0	306	2	C71498	probable yop trans
35	16	50.0	317	2	T00500	hypothetical prote
36	16	50.0	321	2	A84792	paired-box-contain
37	16	50.0	349	2	JR0202	hypothetical prote
38	16	50.0	380	2	T04007	hypothetical prote
39	16	50.0	397	2	T25077	hypothetical prote
40	16	50.0	422	2	S41514	RAU52 protein homo
41	16	50.0	473	2	C81039	lipopolysaccharide
42	16	50.0	473	2	C81984	probable lipopolys
43	16	50.0	475	2	T36342	probable glutamate
44	16	50.0	486	2	B86460	hypothetical prote
45	16	50.0	510	2	A55207	glycerol-3-phospha
46	16	50.0	537	2	G31277	guanine transport
47	16	50.0	580	2	B38418	jockey protein 1 -
48	16	50.0	638	2	T22518	hypothetical prote
49	16	50.0	642	2	C81345	hypothetical prote
50	16	50.0	648	2	T37581	probable serine-ri
51	16	50.0	708	1	TFBOL	lactotransferrin p
52	16	50.0	708	2	UC3323	lactoferrin - goat
53	16	50.0	711	1	TFHUL	lactotransferrin p
54	16	50.0	907	2	A86460	99.9K hypothetical
55	16	50.0	975	2	T29908	hypothetical prote
56	16	50.0	1031	2	T06130	hypothetical prote
57	16	50.0	1069	2	T22138	hypothetical prote
58	16	50.0	1105	2	T22132	hypothetical prote
59	16	50.0	1136	2	AB1581	different protein
60	16	50.0	1380	2	T18309	receptor-adenylate
61	16	50.0	1474	2	D88550	protein ZC84.6 [lm
62	16	50.0	2233	2	T28669	surface protein 51
63	16	50.0	2395	1	S50820	surface protein ty
64	16	50.0	2533	2	T28675	alpha-51D immobili
65	16	50.0	2533	2	T28674	alpha-51D-immobil
66	16	50.0	2543	2	T31687	surface antigen - P
67	16	50.0	2718	2	A23475	G surface protein
68	16	46.9	40	1	SMRF	metallothionein Mt
69	15	46.9	40	2	B61194	metallothionein Mt
70	15	46.9	43	2	S33483	hypothetical prote
71	15	46.9	74	2	A82210	hypothetical prote
72	15	46.9	93	2	T18118	hypothetical prote
73	15	46.9	109	2	S64309	hypothetical prote
74	15	46.9	111	2	A72079	hypothetical prote
75	15	46.9	111	2	B86544	hypothetical prote
76	15	46.9	113	2	T10136	hypothetical prote
77	15	46.9	123	2	B81120	probable periplasm
78	15	46.9	124	2	C81707	hypothetical prote
79	15	46.9	131	2	T36866	probable iron-sulf
80	15	46.9	135	2	T38956	very hypothetical
81	15	46.9	139	2	D97348	hypothetical prote
82	15	46.9	141	2	A80029	Merr-family transc
83	15	46.9	141	2	PC1294	trichophorolite surfac
84	15	46.9	146	2	T01683	immobilization ant
85	15	46.9	149	2	AF2977	conserved hypothet
86	15	46.9	158	2	T14975	hypothetical prote
87	15	46.9	160	2	F98305	hypothetical prote
88	15	46.9	169	1	S18946	ultra high-sulfur
89	15	46.9	171	2	T10315	hypothetical prote
90	15	46.9	179	2	T22471	hypothetical prote
91	15	46.9	181	2	T31481	hypothetical prote
92	15	46.9	185	2	T51003	hypothetical prote
93	15	46.9	189	2	B90969	probable tail asse
94	15	46.9	191	2	S69735	hypothetical prote
95	15	46.9	195	2	T19617	hypothetical prote
96	15	46.9	199	2	C90734	probable outer mem
97	15	46.9	199	2	B85584	hypothetical prote
98	15	46.9	203	2	H81695	recombination prot
99	15	46.9	204	2	A28864	proteinase B (EC 3
100	15	46.9	207	2	A91177	Shut-like protein
101	15	46.9	207	2	B86023	hypothetical prote
102	15	46.9	208	2	T07732	tuberculosis-relat

103	15	46.9	211	2	S34274	probable secretory	176	15	46.9	459	2	14854	gene murine tumour
104	15	46.9	211	2	UC4883	scylladopepsin B	177	15	46.9	461	1	A3536	tumor necrosis fac
105	15	46.9	212	2	T22463	hypothetical prote	178	15	46.9	470	2	A40697	63k sperm flagella
106	15	46.9	214	2	C70812	probable 1pqi prot	179	15	46.9	471	2	T39571	probable regulator
107	15	46.9	219	1	ORECTC	carboxic dehydrat	180	15	46.9	474	2	B38634	tumor necrosis fac
108	15	46.9	219	2	H90677	carbonic anhydrase	181	15	46.9	474	2	T43504	sp66 protein - fas
109	15	46.9	219	2	D85528	carbonic anhydrase	182	15	46.9	476	2	T29054	G protein-coupled
110	15	46.9	220	2	S22493	coat protein - Cym	183	15	46.9	482	2	S65766	phenophorin III -
111	15	46.9	223	2	T33194	hypothetical prote	184	15	46.9	488	2	T10794	hypothetical prote
112	15	46.9	223	2	T33193	hypothetical prote	185	15	46.9	493	2	S46826	chitinase (EC 3.2.
113	15	46.9	227	2	C90899	probable tail asse	186	15	46.9	493	2	S36931	env polypepten -
114	15	46.9	230	2	T49555	related to high cy	187	15	46.9	505	2	S31493	HSP1 protein - yea
115	15	46.9	231	2	S46511	isopenyl transfer	188	15	46.9	503	2	S59698	probable carnitine
116	15	46.9	231	2	S41894	gene 3 protein	189	15	46.9	504	2	C85485	probable carnitine
117	15	46.9	238	1	UH0367	ribonuclease (EC 3	190	15	46.9	504	2	C90634	probable carnitine
118	15	46.9	242	2	AB2007	transposase all160	191	15	46.9	504	2	H64724	probable carnitine
119	15	46.9	245	1	T31838	hypothetical prote	192	15	46.9	505	2	AD0511	probable carnitine
120	15	46.9	245	1	KYBOA	chymotrypsin (EC 3	193	15	46.9	505	2	T19971	hypothetical prote
121	15	46.9	245	1	KYBOB	chymotrypsin (EC 3	194	15	46.9	506	2	T19973	hypothetical prote
122	15	46.9	247	2	B90997	probable tail asse	195	15	46.9	512	2	S28267	phenophorin I prec
123	15	46.9	247	2	T27778	hypothetical prote	196	15	46.9	514	2	T06585	ammonium transport
124	15	46.9	248	2	T03868	hypothetical prote	197	15	46.9	525	2	F84933	IMP cyclohydrolase
125	15	46.9	248	2	T03869	hypothetical prote	198	15	46.9	526	2	T07082	lycopen epsilon-c
126	15	46.9	248	2	T31841	hypothetical prote	199	15	46.9	529	2	S51477	extracellular glyc
127	15	46.9	256	1	TRSP	trypsin-like prote	200	15	46.9	529	2	T48253	myb-like protein -
128	15	46.9	256	1	T27836	hypothetical prote	201	15	46.9	533	2	T00742	ubiquitin-binding
129	15	46.9	258	2	C86286	protein Pg1.16 (I	202	15	46.9	543	2	S65462	glucose transport
130	15	46.9	260	2	E90344	hypothetical prote	203	15	46.9	545	2	A46281	tetrahydrofolylpol
131	15	46.9	262	1	G64977	hydroxyethylthiazol	204	15	46.9	546	2	T11217	reverse transcript
132	15	46.9	262	2	B85837	hydroxyethylthiazol	205	15	46.9	549	2	TN0553	triacylglycerol 1i
133	15	46.9	262	2	C90992	hydroxyethylthiazol	206	15	46.9	551	2	B88949	protein R09B5.4 (I
134	15	46.9	263	1	KXRTB	chymotrypsin (EC 3	207	15	46.9	553	2	T01479	hypothetical prote
135	15	46.9	263	1	A31299	chymotrypsin (EC 3	208	15	46.9	561	2	T27318	hypothetical prote
136	15	46.9	263	2	A21195	chymotrypsin (EC 3	209	15	46.9	562	1	R8BY52	regulatory protein
137	15	46.9	264	2	T38136	chymotrypsin-like	210	15	46.9	568	2	T31692	hypothetical prote
138	15	46.9	264	2	P96807	unknown protein T3	211	15	46.9	571	2	T52325	pectinesterase (EC
139	15	46.9	273	2	A05113	hypothetical prote	212	15	46.9	576	2	H87414	TPR domain protein
140	15	46.9	274	2	S64523	translacion elonga	213	15	46.9	584	2	S14952	pectinesterase hom
141	15	46.9	293	2	A85817	hypothetical prote	214	15	46.9	588	2	T52330	pectinesterase (EC
142	15	46.9	293	2	B85718	probable tail comp	215	15	46.9	588	2	T02184	probable pectinest
143	15	46.9	293	2	T31840	hypothetical prote	216	15	46.9	595	2	T52327	pectinesterase (EC
144	15	46.9	294	2	B84640	CONSTANS-like B-do	217	15	46.9	595	2	T02183	probable pectinest
145	15	46.9	297	2	H72779	hypothetical prote	218	15	46.9	599	2	T10798	phenophorin-S - Vo
146	15	46.9	307	2	T08125	CONSTANS protein h	219	15	46.9	616	2	T29234	hypothetical prote
147	15	46.9	307	2	T45213	superoxide dismuta	220	15	46.9	626	2	T27319	hypothetical prote
148	15	46.9	311	2	T15268	hypothetical prote	221	15	46.9	634	1	H64390	carbon-monoxide de
149	15	46.9	312	2	JCS522	uncoupling protein	222	15	46.9	638	2	T27959	hypothetical prote
150	15	46.9	312	2	T31834	hypothetical prote	223	15	46.9	641	2	T05497	hypothetical prote
151	15	46.9	323	2	S72915	hypothetical prote	224	15	46.9	646	2	T35002	probable respirato
152	15	46.9	326	2	T19996	hypothetical prote	225	15	46.9	661	2	E71427	hypothetical prote
153	15	46.9	337	2	T23794	hypothetical prote	226	15	46.9	670	2	S36616	regulatory protein
154	15	46.9	344	2	S42383	hypothetical prote	227	15	46.9	676	2	T30480	envelope protein h
155	15	46.9	345	2	A05279	hypothetical prote	228	15	46.9	698	1	TFRHP	transferrin precu
156	15	46.9	363	2	JCS536	surface antigen 51	229	15	46.9	698	2	D90771	transferrin prote
157	15	46.9	368	2	S18758	C-Fringe protein 1	230	15	46.9	698	2	H85633	hypothetical prote
158	15	46.9	379	2	B84073	hypothetical prote	231	15	46.9	698	2	F64839	yeast protein precu
159	15	46.9	382	2	T27058	sulfate adenylyltr	232	15	46.9	700	2	T32205	hypothetical prote
160	15	46.9	385	2	T25492	hypothetical prote	233	15	46.9	703	2	A45543	lactoferrin precu
161	15	46.9	386	2	T32240	hypothetical prote	234	15	46.9	713	2	A35502	major surface-labe
162	15	46.9	390	2	C95954	hypothetical prote	235	15	46.9	716	2	T09462	juvenile hormone r
163	15	46.9	392	2	S65693	hypothetical prote	236	15	46.9	717	2	S12100	transferrin precu
164	15	46.9	395	1	A35629	opioid receptor mu	237	15	46.9	717	2	T25431	hypothetical prote
165	15	46.9	400	2	T36553	mevalonate kinase	238	15	46.9	733	2	A95421	probable oxidoredu
166	15	46.9	410	2	T39115	mu opiate receptor	239	15	46.9	738	2	T26421	hypothetical prote
167	15	46.9	412	2	AF0668	formamidase-like p	240	15	46.9	739	2	S40992	hypothetical prote
168	15	46.9	423	1	TFV228	transforming prote	241	15	46.9	738	2	B88553	protein K04H4.2b (
169	15	46.9	423	2	S38953	carboxypeptidase D	242	15	46.9	744	2	S57061	juvenile hormone
170	15	46.9	423	2	A41204	carboxypeptidase B	243	15	46.9	748	2	B45045	nuclear pore prote
171	15	46.9	450	1	S13730	pmbA protein - Esc	244	15	46.9	749	2	B95889	basic juvenile hor
172	15	46.9	450	2	AE1056	probable PmbA prot	245	15	46.9	750	2	H91074	probable dehydroge
173	15	46.9	450	2	D86121	maturation of anti	246	15	46.9	750	2	G85919	transcription regu
174	15	46.9	450	2	D91280	maturation of anti	247	15	46.9	750	2	D65051	hypr protein - Esc
175	15	46.9	454	2	T34297	hypothetical prote	248	15	46.9	751	2	F87789	protein C34G6.2 (I

249	15	46.9	752	2	G85941	hypothetical prote	322	15	46.9	2090	2	T30075	hypothetical prote
250	15	46.9	752	2	C91096	hypothetical prote	323	15	46.9	2153	2	T30074	hypothetical prote
251	15	46.9	752	2	B65070	probable oxidoredu	324	15	46.9	2201	2	A54774	ATP binding cassel
252	15	46.9	755	2	A44315	carriage oligomer	325	15	46.9	2276	2	T34022	zonadhesin - pig
253	15	46.9	771	1	W2B52	gene 52 protein -	326	15	46.9	2531	2	T31070	norch homolog - se
254	15	46.9	793	1	KXMSF	furin (EC 3.4.21.7	327	15	46.9	2844	2	S28291	hypothetical prote
255	15	46.9	793	1	KXRTF	furin (EC 3.4.21.7	328	15	46.9	2910	2	T42214	otogelin - mouse
256	15	46.9	794	1	KXHPF	furin (EC 3.4.21.7	329	15	46.9	2918	2	A54105	fibillin-2 precur
257	15	46.9	797	1	T46044	furin (EC 3.4.21.7	330	15	46.9	3097	2	T00021	DN-cadherin - frii
258	15	46.9	806	2	T18840	hypothetical prote	331	15	46.9	3147	2	T21328	hypothetical prote
259	15	46.9	814	2	T49207	receptor kinase-11	332	15	46.9	3712	2	S18253	lamann alpha-1 ch
260	15	46.9	846	1	VCLJND	env polypeptide pr	333	15	46.9	4660	2	T42737	gp330 protein prec
261	15	46.9	847	1	S36337	histidine decarbox	334	15	46.9	4957	2	T03455	ALR protein - huma
262	15	46.9	849	2	T49342	fzr related protei	335	15	46.9	5262	2	T03454	sublinosin A [vni
263	15	46.9	852	2	S41886	DNA repair protein	336	14	43.8	43	2	A69704	fulvicoin C - Myoc
264	15	46.9	853	2	S54384	envelope polypep	337	14	43.8	45	1	FRYZ	calitoxin - sea an
265	15	46.9	854	1	VCLJST	env polypeptide pr	338	14	43.8	46	2	A31863	vib7 protein - Ag
266	15	46.9	855	1	VCLJZR	env polypeptide pr	339	14	43.8	55	1	B7AG46	lipoprotein R21 pr
267	15	46.9	856	1	A44963	env polypeptide pr	340	14	43.8	60	1	UN0750	lipoprotein R21 pr
268	15	46.9	863	2	A53034	gag polypeptide -	341	14	43.8	60	2	H90831	metallothionein-2a
269	15	46.9	863	2	T27958	hypothetical prote	342	14	43.8	62	2	S54336	metallothionein-2c
270	15	46.9	873	2	B87049	hypothetical prote	343	14	43.8	62	2	S54335	hypothetical prote
271	15	46.9	888	1	GMLJHD	conserved hypotet	344	14	43.8	62	2	T10302	hypothetical prote
272	15	46.9	890	2	T21000	hypothetical prote	345	14	43.8	65	2	S77379	hypothetical prote
273	15	46.9	891	2	E96590	hypothetical prote	346	14	43.8	73	2	F70641	hypothetical prote
274	15	46.9	895	2	B96775	hypothetical prote	347	14	43.8	73	2	G97142	hypothetical prote
275	15	46.9	911	2	S46497	aspartate kinase (348	14	43.8	75	2	A10863	probable lipoprote
276	15	46.9	911	2	B86438	hypothetical prote	349	14	43.8	76	2	E91087	hypothetical prote
277	15	46.9	914	2	T17233	hypothetical prote	350	14	43.8	76	2	E65063	hypothetical prote
278	15	46.9	915	2	T21773	hypothetical prote	351	14	43.8	77	2	T10276	hypothetical prote
279	15	46.9	928	2	T20035	hypothetical prote	352	14	43.8	78	1	IKEC51	estrogen V immunity
280	15	46.9	922	2	A30363	glycoprotein GP330	353	14	43.8	80	2	PM0044	estrogen receptor
281	15	46.9	975	2	T48107	hypothetical prote	354	14	43.8	83	2	H81188	hypothetical prote
282	15	46.9	978	2	H86319	hypothetical prote	355	14	43.8	85	2	E83008	hypothetical prote
283	15	46.9	982	1	VCLJLK	env polypeptide -	356	14	43.8	86	2	S78488	Ig kappa chain V r
284	15	46.9	995	2	S50358	hypothetical prote	357	14	43.8	86	2	T07829	systemic acquired
285	15	46.9	1012	1	DJB665	DNA-directed DNA p	358	14	43.8	88	2	JC1126	major allergen cha
286	15	46.9	1012	2	T44185	probable DNA-dirc	359	14	43.8	91	2	J80036	Clara cell 10K pro
287	15	46.9	1012	2	T43998	DNA polymerase [m	360	14	43.8	91	2	B87324	hypothetical prote
288	15	46.9	1012	2	T41940	DNA polymerase - h	361	14	43.8	91	2	T37327	insulin homolog ce
289	15	46.9	1013	2	T01920	DNA-directed DNA p	362	14	43.8	92	1	UCGMS	proteinase inhibit
290	15	46.9	1014	2	T30431	DNA-directed DNA p	363	14	43.8	96	2	C71614	uteroglobin precu
291	15	46.9	1018	2	T19693	hypothetical prote	364	14	43.8	96	2	A36581	ribosomal protein
292	15	46.9	1047	2	S19508	MSH3 protein - yea	365	14	43.8	96	2	F70534	polychlorinated bi
293	15	46.9	1059	2	T22545	hypothetical prote	366	14	43.8	97	2	S72866	hypothetical prote
294	15	46.9	1063	2	T46284	hypothetical prote	367	14	43.8	98	2	A10334	conserved hypotet
295	15	46.9	1064	2	A40136	fibropellin Ia - s	368	14	43.8	99	2	UC2136	monocyte chemotac
296	15	46.9	1083	2	S59780	hypothetical prote	369	14	43.8	99	2	S60230	glibetellin-regula
297	15	46.9	1095	2	T13964	probable histone d	370	14	43.8	100	2	S64316	hypothetical prote
298	15	46.9	1108	2	T16875	mucin, tracheobron	371	14	43.8	102	2	H84599	hypothetical prote
299	15	46.9	1136	2	A48292	different proteins	372	14	43.8	103	2	S70182	hypothetical prote
300	15	46.9	1136	2	AH1227	magnesium-protopor	373	14	43.8	103	2	S42212	hydroxymethylgluta
301	15	46.9	1193	2	T50729	DNA-directed RNA p	374	14	43.8	105	2	S42212	hydroxymethylgluta
302	15	46.9	1224	2	T07446	protein T12C24.22	375	14	43.8	105	2	S17345	hydroxymethylgluta
303	15	46.9	1245	2	D86260	transferrin-like p	376	14	43.8	105	2	S42213	hydroxymethylgluta
304	15	46.9	1274	2	T10729	cysteine rich prot	377	14	43.8	105	2	T49527	pectin esterase ho
305	15	46.9	1274	2	T42017	RNA2 polypeptide -	378	14	43.8	105	2	T14406	galactose-binding
306	15	46.9	1324	2	S06187	protein FIN21.4 [l	379	14	43.8	105	2	A37961	zinc finger protei
307	15	46.9	1357	2	B96696	hypothetical prote	380	14	43.8	105	2	A13178	hypothetical prote
308	15	46.9	1391	2	T20406	protein-tyrosine k	381	14	43.8	105	4	S57386	putidaredoxin [val
309	15	46.9	1477	2	T18534	polypeptide - Afri	382	14	43.8	106	4	PKPSEP	hypothetical prote
310	15	46.9	1524	2	T30337	hypothetical prote	383	14	43.8	107	1	T49527	Acotf-5 protein -
311	15	46.9	1544	2	T04464	hypothetical prote	384	14	43.8	107	2	E72850	orfs homolog orf13
312	15	46.9	1584	2	T00026	brain-specific ang	385	14	43.8	109	2	T41895	conserved hypotet
313	15	46.9	1626	2	T26318	hypothetical prote	386	14	43.8	109	2	A13178	hypothetical prote
314	15	46.9	1687	2	T30176	EGF repeat transme	387	14	43.8	111	2	S64475	hypothetical prote
315	15	46.9	1722	2	B89753	protein F1IC7.4 [l	388	14	43.8	112	2	S62929	gonadotropin I bet
316	15	46.9	1743	2	T26859	hypothetical prote	389	14	43.8	113	2	S07216	hypothetical prote
317	15	46.9	1820	2	A55494	latent transforin	390	14	43.8	113	2	T20299	hypothetical prote
318	15	46.9	1827	2	A35694	cutl protein - fls	391	14	43.8	113	2	S15197	hyp4 protein - Esc
319	15	46.9	1828	2	T41455	breast/ovarian can	392	14	43.8	116	1	P85921	pleiotropic eftec
320	15	46.9	1863	1	A58881	genome polypeptide	393	14	43.8	116	2	F91076	hyp4 protein (limpo
321	15	46.9	1866	1	GNME2C		394	14	43.8	116	2		

395	14	43.8	116	2	A54598	universal minicirc
396	14	43.8	116	2	D95280	hypothetical prote
397	14	43.8	116	2	F81285	probable periplasm
398	14	43.8	118	1	PSSNM1	phospholipase A2 (
399	14	43.8	118	1	PSSNM3	phospholipase A2 (
400	14	43.8	118	2	D34860	phospholipase A2 (
401	14	43.8	118	2	E34860	phospholipase A2 (
402	14	43.8	118	2	F34860	phospholipase A2 (
403	14	43.8	118	2	F34860	phospholipase A2 (
404	14	43.8	118	2	H34860	phospholipase A2 (
405	14	43.8	118	2	AB0847	Hypa protein (impo
406	14	43.8	119	2	AC3199	transcription regu
407	14	43.8	120	2	E70424	pilin - Aquifex ae
408	14	43.8	121	2	H71351	probable ribosomal
409	14	43.8	123	2	T04635	hypothetical prote
410	14	43.8	124	2	E70754	hypothetical prote
411	14	43.8	125	2	AF2154	hypothetical prote
412	14	43.8	127	2	F70910	hypothetical prote
413	14	43.8	128	2	D69326	conserved hypotet
414	14	43.8	129	2	AH3212	hypothetical prote
415	14	43.8	131	1	BGE02	spermatid transiti
416	14	43.8	133	2	AB1108	hypothetical prote
417	14	43.8	137	2	B36179	gonadotropin I bet
418	14	43.8	137	2	IS1231	gonadotropin I bet
419	14	43.8	137	2	S34349	gonadotropin I bet
420	14	43.8	137	2	S55364	serine proteinase
421	14	43.8	138	2	T21792	hypothetical prote
422	14	43.8	139	2	S69458	hypothetical prote
423	14	43.8	139	2	PC4217	hypothetical 139 p
424	14	43.8	139	2	T12618	homobox protein H
425	14	43.8	140	2	T04904	hypothetical prote
426	14	43.8	141	2	A85994	probable transcrip
427	14	43.8	141	2	AB1009	probable Zn(II)-re
428	14	43.8	141	2	E91148	zinc (II) responsi
429	14	43.8	141	2	I67892	embryonic abundan
430	14	43.8	141	2	T09251	probable CO-induce
431	14	43.8	142	2	T51316	ORF MSV132 probabl
432	14	43.8	142	2	T28293	hypothetical prote
433	14	43.8	142	2	B84040	hypothetical prote
434	14	43.8	142	2	US0510	fusaric acid resis
435	14	43.8	142	2	T03976	hypothetical prote
436	14	43.8	142	4	S13768	MHC class I histoc
437	14	43.8	143	2	H83935	hypothetical prote
438	14	43.8	143	2	T00293	hypothetical prote
439	14	43.8	143	2	T27692	hypothetical prote
440	14	43.8	144	2	S01391	phospholipase A2 (
441	14	43.8	145	2	T05641	hypothetical prote
442	14	43.8	146	2	T25356	hypothetical prote
443	14	43.8	146	2	C86187	YUPH12.12 (import
444	14	43.8	146	2	AD3427	transposase BME114
445	14	43.8	147	2	S42552	proline-rich prote
446	14	43.8	148	2	T30627	hypothetical prote
447	14	43.8	148	2	T48981	hypothetical prote
448	14	43.8	148	2	B72513	hypothetical prote
449	14	43.8	150	2	S11441	plasma C precursor
450	14	43.8	151	2	E83550	positive regulator
451	14	43.8	151	2	T09256	heat shock protein
452	14	43.8	152	1	KRSHMC	keratin high-sulfu
453	14	43.8	152	2	I47111	high-sulfur wool m
454	14	43.8	152	2	I47109	high-sulfur wool m
455	14	43.8	152	2	I47108	high-sulfur wool m
456	14	43.8	152	2	I47112	high-sulfur wool m
457	14	43.8	152	2	A84303	hypothetical prote
458	14	43.8	152	2	T09253	heat shock protein
459	14	43.8	152	2	D89753	protein F137.2 (l
460	14	43.8	153	2	F75514	conserved hypotet
461	14	43.8	153	2	T32950	hypothetical prote
462	14	43.8	154	1	GOVLAW	gene X protein - h
463	14	43.8	154	1	QOVLCP	gene X protein - h
464	14	43.8	154	1	QOVLKS	gene X protein - h
465	14	43.8	154	2	S47408	gene X protein - h
466	14	43.8	154	2	S20756	gene X protein - h
467	14	43.8	154	2	S67503	gene X protein - h
468	14	43.8	154	2	S12542	gene X protein - h
469	14	43.8	154	2	S33687	gene X protein - h
470	14	43.8	155	2	D72761	hypothetical prote
471	14	43.8	156	2	I67751	dopamine receptor
472	14	43.8	156	2	B83185	probable transcrip
473	14	43.8	156	2	B82073	sigma-E factor reg
474	14	43.8	159	1	I83299	sigma-E factor reg
475	14	43.8	159	2	AP0829	sigma-E factor reg
476	14	43.8	159	2	B85903	sigma-E factor, ne
477	14	43.8	159	2	D91058	sigma-E factor reg
478	14	43.8	159	2	T02999	sigma-E factor reg
479	14	43.8	160	2	T28953	hypothetical prote
480	14	43.8	160	2	JC7606	globin chain al pr
481	14	43.8	161	2	T21405	hypothetical prote
482	14	43.8	162	2	AB7755	protein T21B12.3 (
483	14	43.8	166	2	E71342	probable deoxycytl
484	14	43.8	167	2	D81049	hypothetical prote
485	14	43.8	171	2	S09903	hypothetical prote
486	14	43.8	171	2	S09759	hypothetical prote
487	14	43.8	171	2	H83360	probable acetyltra
488	14	43.8	172	2	B81828	hypothetical integ
489	14	43.8	173	2	AB3450	invasion protein b
490	14	43.8	173	2	AB1923	probable membrane
491	14	43.8	174	1	JQ1625	small hydrophobic
492	14	43.8	176	2	T28762	hypothetical prote
493	14	43.8	177	2	B81085	conserved hypotet
494	14	43.8	177	2	S33166	protective antigen
495	14	43.8	177	2	H95981	conserved hypotet
496	14	43.8	178	2	AB1687	conserved hypotet
497	14	43.8	178	2	S64257	hypothetical prote
498	14	43.8	181	2	S78572	EST3 protein - yea
499	14	43.8	182	2	T17625	hypothetical prote
500	14	43.8	184	2	D83234	hypothetical prote
501	14	43.8	185	2	T29777	hypothetical prote
502	14	43.8	187	2	T03990	ZIK protein precur
503	14	43.8	189	2	T28554	hypothetical prote
504	14	43.8	189	2	D36849	AlzL protein - var
505	14	43.8	189	2	B72165	AlzL protein - var
506	14	43.8	189	2	G66514	hypothetical prote
507	14	43.8	189	2	JK0235	core protein MGC-2
508	14	43.8	191	2	JQ1999	hypothetical 21.5K
509	14	43.8	191	2	T32278	hypothetical prote
510	14	43.8	192	2	E42518	AlzL protein - vac
511	14	43.8	193	2	D83296	hypothetical prote
512	14	43.8	196	2	T26943	hypothetical prote
513	14	43.8	197	2	B85068	hypothetical prote
514	14	43.8	197	2	T26829	hypothetical prote
515	14	43.8	197	2	C88809	protein Y43CSB.1 (
516	14	43.8	197	2	C86391	hypothetical prote
517	14	43.8	198	2	S72709	Leb1170 C3 229 pr
518	14	43.8	198	2	A36646	ORF1 protein - equ
519	14	43.8	200	2	B83642	hypothetical prote
520	14	43.8	201	2	S38652	resolvasse - Klebsi
521	14	43.8	201	2	D82779	hypothetical prote
522	14	43.8	203	2	E70150	hypothetical prote
523	14	43.8	204	2	T02386	hypothetical prote
524	14	43.8	205	2	T02385	hypothetical prote
525	14	43.8	205	2	T04918	hypothetical prote
526	14	43.8	207	2	T22436	hypothetical prote
527	14	43.8	207	2	S70533	bbk2.10 protein pr
528	14	43.8	208	2	B75458	hypothetical prote
529	14	43.8	208	2	P00278	hypothetical prote
530	14	43.8	209	2	C70857	probable lipqA prot
531	14	43.8	211	2	S25938	env protein - huma
532	14	43.8	211	2	T18799	hypothetical prote
533	14	43.8	212	2	S09623	agglutinin isolect
534	14	43.8	213	1	A6KWT2	agglutinin isolect
535	14	43.8	213	1	T27841	hypothetical prote
536	14	43.8	213	2	B82698	conserved hypotet
537	14	43.8	215	2	A60166	hemiferrin - bovin
538	14	43.8	215	2	A71541	hypothetical prote
539	14	43.8	216	1	A39684	hemiferrin - rat
540	14	43.8	216	2	H88102	protein W10G11.13

541	14	43.8	216	2	S60048	chlorophyll a/c-bi	614	14	43.8	273	2	G83403	conserved hypothet
542	14	43.8	217	2	T25618	hypothetical prote	615	14	43.8	274	2	JC4171	tryptase (EC 3.4.2
543	14	43.8	218	2	D69428	hypothetical prote	616	14	43.8	275	2	AE3112	transcription regu
544	14	43.8	219	2	T34373	hypothetical prote	617	14	43.8	276	2	H98174	transcription regu
545	14	43.8	219	2	S25939	env. protein - huma	618	14	43.8	275	2	A33210	tryptase (EC 3.4.2
546	14	43.8	220	2	D83390	carbonate dehydrat	619	14	43.8	275	2	G83522	conserved hypothet
547	14	43.8	222	1	MWVZB4	antithrombin-III h	620	14	43.8	275	2	T29952	hypothetical prote
548	14	43.8	223	2	A65172	hypothetical prote	621	14	43.8	277	2	S71222	xyloglucan endo-1,
549	14	43.8	223	2	F89567	protein T08A9.3 [l	622	14	43.8	280	2	D97185	xyloglucan endo-1,
550	14	43.8	223	2	T37974	probable peroxisom	623	14	43.8	280	2	G01884	LIM protein FHL-1,
551	14	43.8	224	2	T01325	hypothetical prote	624	14	43.8	280	2	F95998	probable sugar nuc
552	14	43.8	225	2	C86563	CT465 hypothetical	625	14	43.8	281	2	AE2248	phosphomethylpyrim
553	14	43.8	225	2	C72062	conserved hypothet	626	14	43.8	281	2	C86221	hypothetical prote
554	14	43.8	226	2	T27843	hypothetical prote	627	14	43.8	282	2	T15556	hypothetical prote
555	14	43.8	226	2	T20219	hypothetical prote	628	14	43.8	282	2	T15640	hypothetical prote
556	14	43.8	229	2	T27840	hypothetical prote	629	14	43.8	283	2	T38969	5'-methylthiodeno
557	14	43.8	229	2	B82544	hypothetical prote	630	14	43.8	283	2	T21622	hypothetical prote
558	14	43.8	231	2	S28703	hypothetical prote	631	14	43.8	284	2	H97220	transcriptional reg
559	14	43.8	231	2	F64642	hypothetical prote	632	14	43.8	284	2	T06200	xyloglucan endo-1,
560	14	43.8	231	2	T24669	hypothetical prote	633	14	43.8	284	2	S04278	hypoxanthine phosp
561	14	43.8	232	2	E71325	probable V-type AT	634	14	43.8	285	1	IDECRP	replication initia
562	14	43.8	234	2	T32843	hypothetical prote	635	14	43.8	285	1	I64780	replication-associ
563	14	43.8	237	2	T28070	hypothetical prote	636	14	43.8	285	2	T00306	replication initia
564	14	43.8	238	2	T04166	chamaetin-like pro	637	14	43.8	285	2	B88546	protein R107.2 [im
565	14	43.8	238	2	T40568	hypothetical prote	638	14	43.8	285	2	T24665	hypoxanthine phosp
566	14	43.8	239	2	B81345	probable periplasm	639	14	43.8	285	2	S09614	hypoxanthine phosp
567	14	43.8	240	2	T20319	hypothetical prote	640	14	43.8	286	2	S30872	hypothetical prote
568	14	43.8	242	2	T29854	hypothetical prote	641	14	43.8	286	2	B36570	ethanolamine ammon
569	14	43.8	243	2	A33329	hypothetical prote	642	14	43.8	286	2	T23354	hypothetical prote
570	14	43.8	243	2	JB0204	testis-specific pr	643	14	43.8	286	2	T31494	hypothetical prote
571	14	43.8	244	1	JC2018	testicular protein	644	14	43.8	287	2	A85059	hypothetical prote
572	14	43.8	244	1	T37493	phosducin - mouse	645	14	43.8	288	2	T02546	hypothetical prote
573	14	43.8	245	2	S52096	phosducin oxidore	646	14	43.8	288	2	T21790	hypothetical prote
574	14	43.8	245	2	AF0412	phosducin - cat	647	14	43.8	289	2	T34688	probable lipoprote
575	14	43.8	245	2	T24565	sugar fermentation	648	14	43.8	290	2	C71321	hypothetical prote
576	14	43.8	246	1	A35422	phosducin, retinal	649	14	43.8	290	2	T10104	maturase-like prot
577	14	43.8	247	2	T43324	synaptogyrin homol	650	14	43.8	292	2	T19669	hypothetical prote
578	14	43.8	248	2	S49323	chymotrypsin (EC 3	651	14	43.8	292	2	G88783	protein C33A12.16
579	14	43.8	248	2	T21786	hypothetical prote	652	14	43.8	293	2	E49539	xyloglucan endo-1,
580	14	43.8	248	2	T23769	hypothetical prote	653	14	43.8	294	2	A48844	TGF alpha-like pro
581	14	43.8	249	2	T31837	hypothetical prote	654	14	43.8	295	2	G65018	ethanolamine ammon
582	14	43.8	249	2	T21785	hypothetical prote	655	14	43.8	295	2	B85887	ethanolamine ammon
583	14	43.8	250	2	T31836	hypothetical prote	656	14	43.8	295	2	G91042	ethanolamine ammon
584	14	43.8	250	2	T31835	hypothetical prote	657	14	43.8	295	2	C90740	probable transmemb
585	14	43.8	251	2	T14407	pectin esterase ho	658	14	43.8	295	2	A46017	zinc finger protei
586	14	43.8	251	2	A55523	hypothetical prote	659	14	43.8	295	2	E85590	probable transmemb
587	14	43.8	251	2	B39685	GRESAG protein 2.1	660	14	43.8	295	2	B64818	probable membrane
588	14	43.8	251	2	AB0213	conserved hypothet	661	14	43.8	296	2	H87511	hypothetical prote
589	14	43.8	252	2	S61178	hypothetical prote	662	14	43.8	297	2	T18660	hypothetical prote
590	14	43.8	253	2	T04059	hypothetical prote	663	14	43.8	297	2	A98294	hypothetical prote
591	14	43.8	253	2	E89130	protein F52E1.2 [i	664	14	43.8	297	2	AH2989	permease (imported
592	14	43.8	254	2	AB1304	dihydroorotate deh	665	14	43.8	297	2	S06267	surface antigen H
593	14	43.8	254	2	AB1676	dihydroorotate deh	666	14	43.8	298	2	AF0813	ethanolamine ammon
594	14	43.8	256	2	F82304	sugar fermentation	667	14	43.8	299	2	T43989	minor capsid prote
595	14	43.8	257	2	T28946	hypothetical prote	668	14	43.8	300	2	C88970	transcription regu
596	14	43.8	257	2	T47491	hypothetical prote	669	14	43.8	301	2	F82287	transcription regu
597	14	43.8	258	2	B82249	uridine phosphoryl	670	14	43.8	301	2	B84325	tRNA-pseudouridine
598	14	43.8	259	1	IOHUI	insulin-like growt	671	14	43.8	302	2	D65054	hypothetical prote
599	14	43.8	259	2	T27783	hypothetical prote	672	14	43.8	302	2	H84731	hypothetical prote
600	14	43.8	259	2	E84427	hypothetical prote	673	14	43.8	302	2	JC2071	chitinase (EC 3.2.
601	14	43.8	262	2	T05084	hypothetical prote	674	14	43.8	304	2	A72596	hypothetical prote
602	14	43.8	263	2	S44668	hypothetical prote	675	14	43.8	305	2	T28005	hypothetical prote
603	14	43.8	264	2	T09324	ZK370.7 protein -	676	14	43.8	305	2	S32834	methylviologen-red
604	14	43.8	264	2	AE2274	capsid protein - h	677	14	43.8	310	2	T41985	hypothetical prote
605	14	43.8	265	2	S62363	D-2-chloropropionl	678	14	43.8	311	2	A28446	transferrin - mous
606	14	43.8	267	2	C81651	conserved hypothet	679	14	43.8	311	2	JC7873	L-Thamose-binding
607	14	43.8	267	2	H86548	polymorphic outer	680	14	43.8	312	2	T09443	zinc metalloprotei
608	14	43.8	270	2	I37278	complement factor	681	14	43.8	312	2	E82276	hypothetical prote
609	14	43.8	270	2	C96731	unknown protein F5	682	14	43.8	312	2	T17969	hypothetical prote
610	14	43.8	270	2	F89632	protein F13E6.3 [i	683	14	43.8	312	2	T45061	hypothetical prote
611	14	43.8	271	2	T40238	hypothetical prote	684	14	43.8	314	2	T28879	hypothetical prote
612	14	43.8	272	2	T15351	hypothetical prote	685	14	43.8	317	2	A33985	wound-inducible ch
613	14	43.8	272	2	T21789	hypothetical prote	686	14	43.8	318	2	T04403	probable chitinase

687	14	43.8	319	2	H83253	conserved hypotnet	760	14	43.8	369	2	S60186	peridin-chloroph
688	14	43.8	319	2	B86395	probable origin re	761	14	43.8	370	2	S15013	wnt-1 protein - ze
689	14	43.8	320	2	S38670	chitinase (EC 3.2.	762	14	43.8	372	1	W2ML31	E2 protein - human
690	14	43.8	320	2	T28992	hypothetical prote	763	14	43.8	372	2	T31060	hypothetical prote
691	14	43.8	322	1	A29928	membrane-associat	764	14	43.8	372	2	T45410	hypothetical prote
692	14	43.8	322	2	H87131	probable transcrip	765	14	43.8	372	2	T29359	hypothetical prote
693	14	43.8	323	2	G87358	dieneolactone hydr	766	14	43.8	373	2	A70856	probable lppz prot
694	14	43.8	323	2	T32279	hypothetical prote	767	14	43.8	375	2	T17488	hypothetical prote
695	14	43.8	324	2	A47172	transforming growt	768	14	43.8	375	2	G02300	pancreatic polypep
696	14	43.8	324	2	JC2395	fas antigen precur	769	14	43.8	375	2	S19182	neuropeptide Y/pep
697	14	43.8	325	2	I65354	stem cell antigen	770	14	43.8	375	2	S63685	neuropeptide Y/pep
698	14	43.8	327	2	P96814	hypothetical prote	771	14	43.8	375	2	S52143	amc protein - Brw
699	14	43.8	328	2	G89152	protein C24B5.5 [i	772	14	43.8	375	2	A96720	hypothetical prote
700	14	43.8	328	2	T22636	hypothetical prote	773	14	43.8	378	2	B64044	rod shape-determ
701	14	43.8	329	2	T18619	hypothetical prote	774	14	43.8	379	2	G82380	hypothetical prote
702	14	43.8	331	2	T06227	peroxidase (EC 1.1	775	14	43.8	380	2	C96754	Similar to part of
703	14	43.8	331	2	C71468	probable glycerol-	776	14	43.8	381	2	G83835	sulfate adenylyl tr
704	14	43.8	331	2	B83486	cobalamin biosynth	777	14	43.8	382	2	T27076	hypothetical prote
705	14	43.8	332	2	T04484	probable chitinase	778	14	43.8	382	2	B46133	neuropeptide Y/pep
706	14	43.8	332	2	T21458	hypothetical prote	779	14	43.8	382	2	S27388	neuropeptide Y/pep
707	14	43.8	332	2	A23663	hypothetical prote	780	14	43.8	382	2	S27388	stem cell antigen
708	14	43.8	333	2	D88970	protein f15B1.11	781	14	43.8	382	2	B91083	hypothetical membr
709	14	43.8	333	2	T34498	hypothetical prote	782	14	43.8	382	2	P85928	hypothetical prote
710	14	43.8	334	2	B81272	probable sugar nuc	783	14	43.8	383	1	VGBEKG	glycoprotein precu
711	14	43.8	336	2	T15727	hypothetical prote	784	14	43.8	384	2	A45490	neuropeptide Y/pep
712	14	43.8	340	2	UC7695	G protein-coupled	785	14	43.8	384	2	S64735	retrovirus-related
713	14	43.8	341	1	MMVZHI	antithrombin-III h	786	14	43.8	384	2	T38544	probable exopolyp
714	14	43.8	341	1	H89980	conserved hypotet	787	14	43.8	385	2	S68780	dopamine D1-like y
715	14	43.8	344	2	E72173	B1R protein - vari	788	14	43.8	385	2	S53718	homeotic protein d
716	14	43.8	344	2	S46871	hypothetical prote	789	14	43.8	385	2	A54785	preadipocyte facto
717	14	43.8	344	2	T28611	hypothetical prote	790	14	43.8	386	2	S72168	dopamine receptor
718	14	43.8	344	2	A70771	hypothetical prote	791	14	43.8	386	2	T12527	hypothetical prote
719	14	43.8	345	1	MMVZW2	antithrombin-III h	792	14	43.8	386	2	T19935	hypothetical prote
720	14	43.8	345	1	JN0465	apolipoprotein H p	793	14	43.8	387	2	B71611	hypothetical prote
721	14	43.8	345	2	T32018	hypothetical prote	794	14	43.8	388	2	S51316	prostaglandin E re
722	14	43.8	346	2	AP1820	siatoglycoproteina	795	14	43.8	388	2	I38750	hypothetical 43.6K
723	14	43.8	346	2	T45748	hypothetical prote	796	14	43.8	390	2	S56560	hypothetical prote
724	14	43.8	346	2	T47411	hypothetical prote	797	14	43.8	390	2	B91291	hypothetical prote
725	14	43.8	347	2	S60428	hypothetical prote	798	14	43.8	390	2	D86132	hypothetical prote
726	14	43.8	347	2	JQ2359	wheat aluminum ind	799	14	43.8	392	2	A60777	keratin 2, type I,
727	14	43.8	348	2	S64041	probable membrane	800	14	43.8	392	2	T33505	hypothetical prote
728	14	43.8	348	2	S75548	siatoglycoproteina	801	14	43.8	395	2	S58157	hypothetical prote
729	14	43.8	348	2	S19873	single-stranded DN	802	14	43.8	396	1	A42919	mevalonate kinase
730	14	43.8	349	2	AP3401	enoyl-CoA hydratase	803	14	43.8	397	2	T08345	hypothetical prote
731	14	43.8	349	2	S12863	G protein-coupled	804	14	43.8	398	2	AH0482	conserved hypotet
732	14	43.8	352	2	C96643	hypothetical prote	805	14	43.8	398	2	T19421	hypothetical prote
733	14	43.8	352	2	S77448	hypothetical prote	806	14	43.8	400	2	S32804	beta-3-adrenergic
734	14	43.8	354	2	T26599	hypothetical prote	807	14	43.8	400	2	A41679	beta-3-adrenergic
735	14	43.8	354	2	G86276	hypothetical prote	808	14	43.8	400	2	A53281	beta-3-adrenergic
736	14	43.8	356	2	A25918	thrombomodulin - b	809	14	43.8	400	2	D89810	conserved hypotet
737	14	43.8	356	2	T18590	hypothetical prote	810	14	43.8	401	2	B96570	hypothetical prote
738	14	43.8	357	2	A71295	probable alanine r	811	14	43.8	402	2	T39591	cyclin - fission y
739	14	43.8	357	2	A97697	hypothetical prote	812	14	43.8	403	2	C82228	probable trypsin V
740	14	43.8	358	2	D84901	hypothetical prote	813	14	43.8	403	2	A31847	hypothetical prote
741	14	43.8	358	2	T23802	hypothetical prote	814	14	43.8	403	2	T49480	hypothetical prote
742	14	43.8	359	2	AD3520	galactoside transp	815	14	43.8	405	2	S65459	beta-3-adrenergic
743	14	43.8	361	2	T25305	hypothetical prote	816	14	43.8	405	2	B35340	alpha(1,3)-fucosyl
744	14	43.8	362	1	C65132	hypothetical 38.6	817	14	43.8	406	2	A43850	glutamate N-acetyl
745	14	43.8	362	1	LPECPA	rare lipoprotein A	818	14	43.8	406	2	C81017	amino-acid N-acety
746	14	43.8	362	2	G85563	a minor lipoprotei	819	14	43.8	407	2	T19895	hypothetical prote
747	14	43.8	362	2	G90712	hypothetical prote	820	14	43.8	408	1	QRHUB8	beta-3-adrenergic
748	14	43.8	362	2	T20222	hypothetical prote	821	14	43.8	408	1	QRHUB8	immediate-early pr
749	14	43.8	363	2	I50475	dopamine D1 recep	822	14	43.8	408	2	B83333	probable MPS trans
750	14	43.8	363	2	A34959	triacylglycerol 11	823	14	43.8	408	2	A72869	early gene transac
751	14	43.8	363	2	A81134	acyl-CoA dehydroge	824	14	43.8	410	2	B84144	hypothetical prote
752	14	43.8	363	2	C81888	probable acyl-CoA	825	14	43.8	412	2	B81176	hypothetical prote
753	14	43.8	364	2	JC2115	prostaglandin E re	826	14	43.8	412	2	T26056	hypothetical prote
754	14	43.8	364	2	S65009	collagen alpha 2(I	827	14	43.8	413	1	QRHUB2	beta-2-adrenergic
755	14	43.8	365	2	S10847	neuropeptide Y/pep	828	14	43.8	413	2	S47527	extracellular sucr
756	14	43.8	366	1	S71152	GTP-binding protei	829	14	43.8	413	2	UC2520	beta-fructofuranos
757	14	43.8	367	1	I51426	peccate lyase (EC	830	14	43.8	414	1	QRHUB3	beta-3-adrenergic
758	14	43.8	367	2	S15716	cinnamoyl CoA redu	831	14	43.8	415	2	I53040	coenzyme F420 hydr
759	14	43.8	368	2	T48643		832	14	43.8	415	2	A64299	

979 14 43.8 510 1 S15620 l1 protein - human
980 14 43.8 510 1 S15627 l1 protein - human
981 14 43.8 511 2 S44275 dopamine receptor
982 14 43.8 511 2 JC7682 spermatogenesis as
983 14 43.8 513 2 T38866 conserved hypoch
984 14 43.8 515 2 D71319 probable glucose-6
985 14 43.8 516 2 T33269 hypothetical prote
986 14 43.8 516 2 S40996 hypothetical prote
987 14 43.8 517 2 A71707 lipopolysaccharide
988 14 43.8 519 2 S78089 G-protein signalin
989 14 43.8 520 2 T23545 hypothetical prote
990 14 43.8 522 2 S71821 probable interleuk
991 14 43.8 523 1 S61713 carboxypeptidase C
992 14 43.8 524 1 P1WL58 l1 protein - human
993 14 43.8 527 1 SAHUP saposin precursor
994 14 43.8 528 2 B88989 protein F02C9.2 [1
995 14 43.8 528 2 T33527 hypothetical prote
996 14 43.8 529 2 S46116 probable regulator
997 14 43.8 530 2 UC5939 estrogen receptor
998 14 43.8 531 2 T32276 hypothetical prote
999 14 43.8 532 2 S18448 variant surface g1
1000 14 43.8 533 2 T31002 hypothetical prote

ALIGNMENTS

RESULT 1

S59448
hypothetical protein YMR206w - yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein YMR325.07
C/Species: Saccharomyces cerevisiae
C/Date: 30-Nov-1995 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C/Accession: S59448
R/Odell, C.; Bowman, S.
submitted to the EMBL Data Library, March 1995
A/Reference number: S59441
A/Accession: S59448
A/Molecule type: DNA
A/Residues: 1-313 <ODE>
A/Cross-references: UNIPROT:Q03695; UNIPARC:UPI000013B90A; EMBL:Z48755; NID:G736296; PII
C/Genetics:
A/Gene: MIPS:YMR206w
A/Cross-references: SGD:S0004819
A/Map position: 13R
Query Match 56.2%; Score 18; DB 2; Length 313;
Best Local Similarity 20.0%; Pred. No. 0.44; 8; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXC 13
DB 245 CSSSSSSSSAC 254
RESULT 2
E75433
hypothetical protein - Deinococcus radiodurans (strain R1)
C/Species: Deinococcus radiodurans
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: E75433
R/White, O.; Eilen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uteerback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A/Reference number: A75250; MUID:20036896; PMID:10567266
A/Accession: E75433
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-480 <WHI>
A/Cross-references: UNIPROT:Q9RV87; UNIPARC:UPI00000C18A8; GB:AE001963; GB:AE000513; NID

A/Experimental source: strain R1

C/Genetics:

A/Gene: DR1142

A/Map position: 1

C/Superfamily: Deinococcus radiodurans hypothetical protein DR1142

Query Match 56.2%; Score 18; DB 2; Length 480;
Best Local Similarity 20.0%; Pred. No. 0.51; 8; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 464 CSAATATAC 473

RESULT 3

JQ1560
hypothetical 20.6K protein - Lymantria dispar nuclear polyhedrosis virus
N/Alternate names: hypothetical protein 4
C/Species: Lymantria dispar nuclear polyhedrosis virus, LdNPV
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C/Accession: JQ1560
R./Bjornson, R.M.; Rohmann, G.F.
J. Gen. Virol. 73, 1499-1504, 1992
A/Title: Nucleotide sequence of the polyhedron envelope protein gene region of the Lyman
A/Reference number: PQ0339; MUID:92300345; PMID:1607868
A/Accession: JQ1560
A/Molecule type: DNA
A/Residues: 1-194 <BJO>
A/Cross-references: UNIPROT:P36868; UNIPARC:UPI000017A7BF; DBJ:DI0836

Query Match 53.1%; Score 17; DB 2; Length 194;
Best Local Similarity 20.0%; Pred. No. 1.6; 8; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 74 CSAANTSSSC 83

RESULT 4

JC4082
coat protein - Cymbidium mosaic virus
C/Species: Cymbidium mosaic virus
C/Date: 21-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C/Accession: JC4082
R./Ryu, K.H.; Yoon, K.E.; Park, W.M.
Gene 156, 303-304, 1995
A/Title: Cloning and sequencing of a cDNA encoding the coat protein of a Korean isolate
A/Reference number: JC4082; MUID:95278762; PMID:7758973
A/Accession: JC4082
A/Molecule type: mRNA
A/Residues: 1-220 <RYU>
A/Cross-references: UNIPROT:Q66152; UNIPARC:UPI0000F22DE; EMBL:X81051; NID:G897718; PII
C/Note: The authors translated the codon GTG for residue 161 as Leu
C/Superfamily: potato virus coat protein
C/Keywords: coat protein

Query Match 53.1%; Score 17; DB 2; Length 220;
Best Local Similarity 20.0%; Pred. No. 1.7; 8; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 166 CSAATLTATC 175

RESULT 5

AB2920
cobalamin biosynthetic protein [imported] - Agrobacterium tumefaciens (strain C58, Dupor
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: AB2920

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavina, T.; Levy, R.; Li, M.; McCell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, B.W.
 A>Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AB2920
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-376 <KOR>
 A:Cross-references: UNIPROT:Q8UBQ6; UNIPARC:UPI0000164772; GB:AE008688; P1DN:AAI43776.1;
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: cbid
 A:Map position: circular chromosome
 C:Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 53.1%; Score 17; DB 2; Length 376;
 Best Local Similarity 20.0%; Pred. No. 2;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 17 CAAATKAC 26

RESULT 6
 C97694
 cobalamin biosynthetic protein cbid (PA2908) [imported] - *Agrobacterium tumefaciens* (str
 C:Species: *Agrobacterium tumefaciens*
 C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
 C:Accession: C97694
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2333-2328, 2001
 A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: C97694
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-388 <KUR>
 A:Cross-references: UNIPROT:Q8UBQ6; UNIPARC:UPI000002014; GB:AE007869; P1DN:AAK8508.1;
 C:Genetics:
 A:Gene: AGR_C_5073
 A:Map position: circular chromosome
 C:Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 53.1%; Score 17; DB 2; Length 388;
 Best Local Similarity 20.0%; Pred. No. 2;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 29 CAAATKAC 38

RESULT 7
 T23167
 hypothetical protein K01C8.2 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T23167
 R:Sim, M.
 submitted to the EMBL Data Library, April 1995
 A:Reference number: Z19702
 A:Accession: T23167
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-389 <MTL>
 A:Cross-references: UNIPROT:Q21081; UNIPARC:UPI0000080155; EMBL:Z49066; P1DN:CAA8855.1;
 A:Experimental source: clone K01C8

C:Genetics:
 A:Gene: CESP:K01C8.2
 A:Map position: 2
 A:Introns: 54/2; 146/3; 208/3; 283/1; 379/3

Query Match 53.1%; Score 17; DB 2; Length 389;
 Best Local Similarity 20.0%; Pred. No. 2;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 123 CSSSISTSSC 132

RESULT 8
 T37314
 probable kexin (EC 3.4.21.61) - *Caenorhabditis elegans* (fragment)
 N/Alternate names: Diesterase 4
 C:Species: *Caenorhabditis elegans*
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
 C:Accession: T37314
 R:Thacker, C.; Peters, K.; Strayko, M.; Rose, A.M.
 Genes Dev. 9, 956-971, 1995
 A>Title: The bli-4 locus of *Caenorhabditis elegans* encodes structurally distinct kex2/en
 A:Reference number: Z21679; MUID:95293228; PMID:7774813
 A:Accession: T37314
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-570 <THA>
 A:Cross-references: UNIPARC:UPI000016B8E1; EMBL:IL29440; NID:G459702; P1DN:AAA98752.1; P
 C:Genetics:
 A:Gene: bli-4
 A:Map position: I
 C:Keyword: alternative splicing; hydrolase; serine proteinase

Query Match 53.1%; Score 17; DB 2; Length 570;
 Best Local Similarity 20.0%; Pred. No. 2.3;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 311 CTSSSANTSC 320

RESULT 9
 S33575
 dnaK-type molecular chaperone precursor, mitochondrial - *Leishmania major*
 N/Alternate names: heat shock protein 70-related protein; mitochondrial stress protein
 C:Species: *Leishmania major*
 C>Date: 09-Dec-1997 #sequence_revision 09-Dec-1997 #text_change 31-Dec-2004
 C:Accession: S33575; S78090; S05438
 R:Searle, S.; McCrossan, M.V.; Smith, D.F.
 J. Cell Sci. 104, 1091-1100, 1993
 A>Title: Expression of a mitochondrial stress protein in the protozoan parasite *Leishman*
 A:Reference number: S33575; MUID:93300981; PMID:8314893
 A:Accession: S33575
 A:Molecule type: DNA
 A:Residues: 1-634 <SEA>
 A:Cross-references: UNIPROT:P12076; UNIPARC:UPI0000177D24; EMBL:X64137
 R:Smith, D.F.
 submitted to the EMBL Data Library, January 1992
 A:Reference number: S78090
 A:Accession: S78090
 A:Molecule type: DNA
 A:Residues: 1-460,481-500, 'OGERRASRNQIRGRFSLG', 501-634 <SMI>
 A:Cross-references: UNIPARC:UPI000012C5C; EMBL:X64137; NID:G311289; P1DN:CAA45498.1; E
 R:Searle, S.; Campos, A.J.R.; Coulson, R.M.R.; Spithill, T.W.; Smith, D.F.
 Nucleic Acids Res. 17, 5081-5095, 1989
 A>Title: A family of heat shock protein 70-related genes are expressed in the promastig
 A:Reference number: S05438; MUID:89345072; PMID:2762121
 A:Accession: S05438
 A>Status: translation not shown
 A:Molecule type: DNA

A:Residues: 1-249 <SEF>
A:Cross-references: UNIPARC:UPI000016BPLC; EMBL:X14574; NID:g9561; PIDN:CAA32713.1; PID:
C:Genetics:
A:Gene: lmbp70.1
A:Genome: nuclear
C:Function:
A:Description: involved in protein folding and assembling/disassembling of protein comp
C:Superfamily: bcr protein
C:Keywords: ATP; mitochondrion; molecular chaperone; stress-induced protein
F.1-23/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F.24-634/Product: heat shock protein 70, mitochondrial #status predicted <MAT>

Query Match 53.1%; Score 17; DB 2; Length 634;
Best Local Similarity 20.0%; Pred. No. 2.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 7 CGSAAASAA 16

RESULT 10
D87803
protein bli-4D [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 21-Jun-2002
C:Accession: D87803
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: D87803
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-942 <STO>
A:Cross-references: UNIPARC:UPI000016B639; GB:chr_I; PIDN:AAB96754.1; PID:g2773243; GSPI
C:Genetics:
A:Gene: bli-4D
A:Map position: 1
C:Superfamily: kexin; subtilisin homology

Query Match 53.1%; Score 17; DB 2; Length 942;
Best Local Similarity 20.0%; Pred. No. 2.8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 683 CTSSSATSC 692

RESULT 11
S09118
G surface protein 168 - Paramoecium primaurelia
C:Species: Paramoecium primaurelia
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C:Accession: S09118
R:Prat, A.
J. Mol. Biol. 211, 521-535, 1990
A:Title: Conserved sequences flank variable tandem repeats in two alleles of the G surf
A:Reference number: S09118; MUID:90172419; PMID:2308165
A:Accession: S09118
A:Molecule type: DNA
A:Residues: 1-2704 <PRA>
A:Cross-references: UNIPROT:P17053; UNIPARC:UPI000012AE23; EMBL:X52133; NID:g10049; PIDN
C:Genetics:
A:Genetic code: SGCS
C:Superfamily: G surface protein

Query Match 53.1%; Score 17; DB 2; Length 2704;
Best Local Similarity 20.0%; Pred. No. 4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 1632 CVAATTAATTC 1641

RESULT 12
F84312
hypothetical protein Vng1598h [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: F84312
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laeky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danison, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: F84312
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-71 <STO>
A:Cross-references: UNIPROT:Q9HPL; UNIPARC:UPI0000063920; GB:AE004437; NID:g10581076; P
C:Genetics:
A:Gene: VNG1598H

Query Match 50.0%; Score 16; DB 2; Length 71;
Best Local Similarity 20.0%; Pred. No. 4.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 30 CSRSVSTSTC 39

RESULT 13
E70531
hypothetical protein RV2706c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: E70531
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentile, S.; Hamlin, N.; Holtrold, S
; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: E70531
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-85 <COL>
A:Cross-references: UNIPROT:O07207; UNIPARC:UPI00000C14DC; GB:296072; GB:AL123456; NID:g
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV2706c

Query Match 50.0%; Score 16; DB 2; Length 85;
Best Local Similarity 20.0%; Pred. No. 5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 24 GSATVAAVC 33

RESULT 14
S51479
drought-induced protein D121 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 15-Jul-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S51479; S43176

A:Goetti, F.; Bertauche, N.; Varianian, N.; Giraudat, J.
Mol. Gen. Genet. 246, 10-18, 1995
A:Title: Abscisic acid-dependent and -independent regulation of gene expression by proglutelin
A:Reference number: S51478; MUID:95124290; PMID:7823904
A:Accession: S51479
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-104 <GOS>
A:Cross-references: UNIPROT:Q39084; UNIPARC:UPI000000BEP21, EMBL:X78585, NID:g469111, PMID:7823904
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994
C:Genetics:
A:Gene: D121
C:Superfamily: late embryogenesis-abundant protein leas

Query Match	50.0%;	Score 16;	DB 2;	Length 104;
Best Local Similarity	20.0%;	Pred. No. 5.4;		
Matches	2;	Conservative	0;	Mismatches 8; Indels 0; Gaps 0;
Oy	4	XXXXXXXXXC	13	
Db	14	CSAAGGSUSC	23	

RESULT 15
 A71249
 hypochlorite protein PH0248 - *Pyrococcus horikoshii*
 C:Species: *Pyrococcus horikoshii*
 C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
 C:Accession: A71249
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
 M.; Ohnuki, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi
 D.A. *Res.* 5, 55-76, 1998
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
 A:Reference number: A71000; MUID:98344137; PMID:9679194
 A:Accession: A71249
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-105 <KAM>
 A:Cross-references: UNIPROT:O57986; UNIPARC:UP10000062D66; GS:AP000001; NID:g3236128; PI
 A:Experimental source: strain OT3
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C:Genetics:
 ;Gene: PH0248

Query Match	Score 16;	DB 2;	Length 105;
Best Local Similarity	50.0%;		
Best Local Similarity	Pred. No. 5.4;		
Matches	2; Conservative	0;	Mismatches 8; Indels 0; Gaps 0;
Qy	4 CXXXXXXXXC 13		
db	35 CFTASASTAC 44		

RESULT 16
 F72549
 hypothetical protein APE1682 - Aeropyrum pernix (strain K1)
 C/Species: Aeropyrum pernix
 C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
 C/Accession: F72549
 R/Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahashi, H.; Takamaya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
 DNA Res. 6, 83-101, 1999
 A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
 A/Reference number: A72450; MUID:99310339; PMID:10382966

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <KAW>
A:Cross-references: UNIPROT:Q9YB5; UNIPARC:UPI000005E048; DDBJ:AP000062; NID:G5105244;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APR1682
C:Superfamily: Aeropyrum pernix hypothetical protein APR1682

Query Match	50.0%;	Score 16;	DB 2;	Length 108;
Best Local Similarity	20.0%;	Pred. No. 5.5;		
Matches	2;	Conservative	0;	Mismatches 8;
			Indels	0;
			Gaps	0;

```
QY      4 CXXXXXXXXX 13
         |
Db      8 CFSASKAAC 17
```

RESULT 17
A23473
chymotrypsin-like proteinase (EC 3.4.21.-) - pig (tentative sequence) (fragments)
N:Alternate names: pancreatic elastase II (misidentification)
S:Species: Sus scrofa domestica (domestic pig)
C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 09-Jul-2004
A:Accession: A23473
R:Vered, M.; Gertler, A.; Buzstein, Y.
Int. J. Pept. Protein Res. 27, 183-190, 1986
A:Reference number: A23473; MUID:86194934; PMID:3634756
A:Molecule type: protein
A:Residues: 1-126 <VER>
A:Cross-references: UNIPROT:Q7M325; UNIPARC:UPI0000175C02
C:Superfamily: trypsin; trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase

Query Match	50.0%;	Score 16;	DB 2;	Length 126;
Best Local Similarity	20.0%;	Pred. No..5.8;		
Matches	2;	Conservative	0;	Mismatches 8;
			Indels	0;
			Gaps	0;

QY	4	XXXXXXXXXX	13
Db	87	CAGASGASSC	96

RESULT 18
T49498
hypochemical protein B1AD6.500 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #next_change 18-Aug-2000
C:Accession: T49498
R:Schulte, U., Aign, V., Hohelsel, J., Brandt, P., Fartmann, B., Holland, R., Nyakatura,
submitted to the Protein Sequence Database, May 2000
;Reference number: Z25022

A,Status: preliminary
A,Molecule type: DNA
A,Residues: 1-129 <SCH>
A,Cross-references: UNIPARC:UPI0000179E1D; EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.500
A,Experimental source: BAC clone B14D6; strain OR74A

C:Superfamily, Neurospora crassa hypothetical protein B14D6.500

Query Match	50.0%;	Score 16;	DB 2;	Length 129;
Best Local Similarity	20.0%;	Pred. No. 5,8;		
Matches	2;	Conservative	0;	Mismatches 8;
				Indels

```
QY      4 CXXXXXXXXC 13
Db      90 CSARASISTC 99
```

RESULT 19
A96746
hypothetical protein T9N14.5 [imported] - Arabidopsis thaliana
C|Species: Arabidopsis thaliana (mouse-ear cross)
C|Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #ext_change 09-Jul-2004
C|Accession: A96746
R|Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.: Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: A96746
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-133 <STO>
 A:Cross-references: UNIPROT:Q9C7T1; UNIPARC:UPI000004A50D; GB:AE005173; NID:G10645374; F
 C:Genetics:
 A:Gene: T9N14.5
 A:Map position: 1

Query Match 50.0%; Score 16; DB 2; Length 133;
 Best Local Similarity 20.0%; Pred. No. 5.9;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
 DB 93 CYSSTSSVSC 102

RESULT 20
 S09762
 hypothetical protein TRL3 precursor - human cytomegalovirus (strain AD169)
 C:Species: human cytomegalovirus, human herpesvirus 5
 A:Note: host Homo sapiens (man)
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
 C:Accession: S09762
 R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohm, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;
 M.; Barrall, B.G.
 Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
 A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
 A:Reference number: S09749; MUID:90269039; PMID:2161319
 A:Accession: S09762
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-147 <CHE>
 A:Cross-references: UNIPROT:P16811; UNIPARC:UPI000012D86F; EMBL:X17403; NID:G59591; PIDN
 A:Note: this sequence was submitted to the EMBL Data Library, December 1989
 C:Superfamily: human cytomegalovirus hypothetical protein TRL3
 C:Keywords: glycoprotein
 F:1-35/Domain: signal sequence #status predicted <SIG>
 F:26-147/Product: hypothetical protein TRL3 #status predicted <MAT>
 F:21,35,54,62,99,110,125,133,144/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match 50.0%; Score 16; DB 2; Length 147;
 Best Local Similarity 20.0%; Pred. No. 6.1;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
 DB 97 CTNTTTVTC 106

RESULT 21
 PNO103
 hypothetical 17k protein - barley stripe mosaic virus
 C:Species: barley stripe mosaic virus, BSMV
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
 C:Accession: PNO103
 R:Kozlov, Y.V.; Afanasiev, B.N.; Rupakov, V.V.; Golova, Y.B.; Kulaeva, O.I.; Dolja, V.V.
 Mol. Biol. (Mosk.) 23, 1080-1090, 1989
 A:Title: The complete nucleotide sequence of barley stripe mosaic virus RNA 3 and its va
 A:Reference number: PNO102; MUID:90066400; PMID:2586501
 A:Accession: PNO103
 A:Molecule type: genomic RNA
 A:Residues: 1-153 <KOZ>

A:Cross-references: UNIPROT:Q07118; UNIPARC:UPI00000F58A5
 C:Genetics:
 A:Map position: segment 3(III)

Query Match 50.0%; Score 16; DB 2; Length 153;
 Best Local Similarity 20.0%; Pred. No. 6.2;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
 DB 10 CGTSTSTVC 19

RESULT 22
 T28088
 hypothetical protein ZK899.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T28088
 R:Kershaw, J.
 submitted to the EMBL Data Library, September 1994
 A:Reference number: Z20468
 A:Accession: T28088
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-161 <WIL>
 A:Cross-references: UNIPROT:Q23659; UNIPARC:UPI000007B9F8; EMBL:Z37140; PIDN:CAA85496.1,
 A:Experimental source: clone ZK899
 C:Genetics:
 A:Gene: CESP:ZK899.1
 A:Map position: X 89/3; 144/2
 A:Introns: 46/3; 89/3; 144/2
 C:Superfamily: *Caenorhabditis elegans* hypothetical protein ZK899.1

Query Match 50.0%; Score 16; DB 2; Length 161;
 Best Local Similarity 20.0%; Pred. No. 6.3;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
 DB 152 CSTSTSTSC 161

RESULT 23
 J01252
 hypothetical 16.7k protein - chrysanthemum virus B
 C:Species: chrysanthemum virus B
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
 C:Accession: J01252
 R:Levey, K.; Zayriev, S.
 J. Gen. Virol. 72, 2333-2337, 1991
 A:Title: Nucleotide sequence and gene organization of the 3'-terminal region of chrysant
 A:Reference number: J01246; MUID:92013948; PMID:1919520
 A:Accession: J01252
 A:Molecule type: genomic RNA
 A:Residues: 1-164 <LRY>
 A:Cross-references: UNIPROT:Q7LZW7; UNIPARC:UPI0000179E53

Query Match 50.0%; Score 16; DB 2; Length 164;
 Best Local Similarity 20.0%; Pred. No. 6.3;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
 DB 80 CSLSSAASC 89

RESULT 24
 T34277
 hypothetical protein F46H5.2 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T34277

R.Nhan, M.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid F46H5.

A:Reference number: Z21498

A:Accession: T34277

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-229 <MHA>

A:Cross-references: UNIPROT:Q20490; UNIPARC:UPI0000179CAC; EMBL:U41543; PIDD:AA37021.1;

A:Experimental source: strain Bristol NZ; clone F46H5

C:Genetics:

A:Gene: CESP:F46H5.2

A:Map position: X

A:Introns: 16/1; 85/2; 151/2

Query Match 50.0%; Score 16; DB 2; Length 229;
 Best Local Similarity 20.0%; Pred. No. 7.1;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 23 CEASASNSC 32

RESULT 25

pectate lyase (EC 4.2.2.2) - fungus (Fusarium solani)

C:Species: Fusarium solani

C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: A45724

R:Gonzalez-Candelas, L.; Kolattukudy, P. E.

J. Bacteriol. 174, 6343-6349, 1992

A:Title: Isolation and analysis of a novel inducible pectate lyase gene from the phytoge

A:Reference number: A45724; MUID:93015682; PMID:1400187

A:Accession: A45724

A:Status: preliminary

A:Molecule type: DNA; protein

A:Residues: 1-242 <CON>

A:Cross-references: UNIPROT:Q04701; UNIPARC:UPI000006C8B5; GB:M94691; NID:g168155; PIDD:

A:Experimental source: isolate T8

C:Keywords: carbon-oxygen lyase

Query Match 50.0%; Score 16; DB 2; Length 242;
 Best Local Similarity 20.0%; Pred. No. 7.3;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 233 CTVGTSTSC 242

RESULT 26

chymotrypsin B - Atlantic cod (fragment)

C:Species: Gadus morhua (Atlantic cod)

C:Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 09-Jul-2004

C:Accession: S72219

R:Leht-Larsen, R.; Asgeirsson, B.; Thorolfsson, M.; Norregard-Madsen, M.; Hojrup, P.

Biochim. Biophys. Acta 1297, 49-56, 1996

A:Title: Structure of chymotrypsin variant B from Atlantic cod, Gadus morhua.

A:Reference number: S72219; MUID:96439045; PMID:8841380

A:Accession: S72219

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-14;15-244 <LET>

A:Cross-references: UNIPROT:Q9PM06; UNIPARC:UPI0000175C00; UNIPARC:UPI0000175C01

C:Superfamily: trypsin; trypsin homology

F:15-237/Domain: trypsin homology <TRY>

Query Match 50.0%; Score 16; DB 2; Length 244;
 Best Local Similarity 20.0%; Pred. No. 7.3;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 181 CAGAGATSC 190

RESULT 27

A55035

cysteine-rich protein CRP1 - earthworm (Enchytraeus buchholzi)

C:Species: Enchytraeus buchholzi

C:Date: 14-Nov-1994 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004

C:Accession: A55035; S45034

R:Willuhn, J.; Schmitt-Wrede, H. P.; Greven, H.; Wunderlich, F.

J. Biol. Chem. 269, 24688-24691, 1994

A:Title: cDNA cloning of a cadmium-inducible mRNA encoding a novel cysteine-rich, non-mc

A:Reference number: A55035; MUID:95014230; PMID:7929141

A:Accession: A55035

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-251 <ML>

A:Cross-references: UNIPROT:Q24774; UNIPARC:UPI000007D243; EMBL:X79344; NID:g488802; PIDD:

C:Superfamily: ultra-high-sulfur keratin

Query Match 50.0%; Score 16; DB 2; Length 251;
 Best Local Similarity 20.0%; Pred. No. 7.4;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 28 CGTSTTAGSC 37

RESULT 28

J01724

EI membrane glycoprotein precursor - canine coronavirus (strain Insavec-1)

N:Alternate names: matrix glycoprotein

C:Species: canine coronavirus

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004

C:Accession: J01724

R:Horzburgh, B. C.; Brierley, I.; Brown, T. D. K.

J. Gen. Virol. 73, 2849-2862, 1992

A:Title: Analysis of a 9.6 kb sequence from the 3' end of canine coronavirus genomic RNA

A:Reference number: P00481; MUID:93057357; PMID:1431811

A:Accession: J01724

A:Molecule type: genomic RNA

A:Residues: 1-262 <HOR>

A:Cross-references: UNIPROT:P36299; UNIPARC:UPI0000138986; DBJ:DJ1096; NID:g406193; PIDD:

C:Genetics:

A:Gene: M

C:Superfamily: coronavirus EI membrane glycoprotein

C:Keywords: glycoprotein; matrix protein; transmembrane protein

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-262/Product: EI membrane glycoprotein #status predicted <MAT>

F:56-72/Domain: transmembrane #status predicted <TM1>

F:85-101/Domain: transmembrane #status predicted <TM2>

F:115-134/Domain: transmembrane #status predicted <TM3>

F:32,55/Binding site: carbohydrate (asn) (covalent) #status predicted

F:33,57/Binding site: carbohydrate (ser) (covalent) #status predicted

F:44/Binding site: carbohydrate (thr) (covalent) #status predicted

Query Match 50.0%; Score 16; DB 1; Length 262;
 Best Local Similarity 20.0%; Pred. No. 7.5;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 21 CAMTSSTSC 30

RESULT 29

S47537

chymotrypsin (EC 3.4.21.1) precursor - Atlantic cod

C:Species: Gadus morhua (Atlantic cod)

C.Date: 26-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
A.Accession: S47537, S43163
R.Guimondodotit, A.; Oskarsen, S.; Eakin, A.E.; Craik, C.S.; Bjarnason, J.B.
Biochim. Biophys. Acta 1219, 211-214, 1994
A.Title: Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.
A.Reference number: S47537, MUID:9436860, PMID:8086467
A.Accession: S47537
A.Molecule type: mRNA
A.Residues: 1-263 <GUD>
A.Cross-references: UNIPROT:P47796, UNIPARC:UPI0000128664, EMBL:X78490, NID:9468750, PID
C:Superfamily: trypsin, trypsin homology
C:Keywords: hydrolase; protein digestion; serine proteinase
F.1-18/Domain: signal sequence #status predicted <Sig>
F.19-263/Product: chymotrypsin #status predicted <Mat>
F.34-256/Domain: trypsin homology <TRY>
F.75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 50.0%; Score 16; DB 2; Length 263;
Best Local Similarity 20.0%; Pred. No. 7.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXC 13
Db 200 CAGAGASSC 209

RESULT 30
A5856
Probable elongation factor yelp [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C.Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 05-Oct-2004
A.Accession: A5856
R.Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimianta, E.; Potamoustis, K.; Apodaca,
Natter 409, 529-533, 2001
A.Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A.Reference number: A58480, MUID:21074935, PMID:11206551
A.Accession: A5856
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-275 <STO>
A.Cross-references: UNIPARC:UPI00000020288, GB:AE005174, NID:912516497, PIDN:AAG57309.1;
A.Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A.Gene: yelp
C:Superfamily: translation elongation factor P

Query Match 50.0%; Score 16; DB 2; Length 275;
Best Local Similarity 20.0%; Pred. No. 7.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXC 13
Db 54 CTSKSANTTC 63

RESULT 31
B64986
Hypothetical 30.9 kD protein in frub 5' region - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C.Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 05-Oct-2004
A.Accession: B64986
R.Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.a.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A.Title: The complete genome sequence of Escherichia coli K-12.
A.Reference number: A64720, MUID:97426617, PMID:9278503
A.Accession: B64986
A.Status: preliminary; nucleic acid sequence not shown; translation not shown
A.Molecule type: DNA
A.Residues: 1-275 <BLAT>
A.Cross-references: UNIPARC:UPI0000167DC1, GB:AE000306, GB:U00096, NID:91788489, PIDN:AA
A.Experimental source: strain K-12, substrain MG1655

C:Genetics:
A.Gene: yelp
C:Superfamily: translation elongation factor P

Query Match 50.0%; Score 16; DB 2; Length 275;
Best Local Similarity 20.0%; Pred. No. 7.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXC 13
Db 54 CTSKSANTTC 63

RESULT 32
G91011
Probable elongation factor [imported] - Escherichia coli (strain O157:H7, substrain RMD
C:Species: Escherichia coli
C.Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 05-Oct-2004
A.Accession: G91011
R.Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A.Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A.Reference number: A9629; MUID:21156231, PMID:11258796
A.Accession: G91011
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-275 <HAY>
A.Cross-references: UNIPARC:UPI00000020288, GB:BA000007, PIDN:BA836486.1, PID:913362532;
A.Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A.Gene: RC93063
C:Superfamily: translation elongation factor P

Query Match 50.0%; Score 16; DB 2; Length 275;
Best Local Similarity 20.0%; Pred. No. 7.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXC 13
Db 54 CTSKSANTTC 63

RESULT 33
T23682
Hypothetical protein M02G9.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
A.Accession: T23682
R.Mathews, L.
submitted to the EMBL Data Library, November 1996
A.Reference number: Z19781
A.Accession: T23682
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-294 <MIL>
A.Cross-references: UNIPROT:Q9XUS0, UNIPARC:UPI0000076A8F, EMBL:281573, PIDN:CAB04626.1;
A.Experimental source: clone M02G9
C:Genetics:
A.Gene: CESP:M02G9.3
A.Map position: 2
A.Introns: 20/3; 76/2; 182/3; 223/3

Query Match 50.0%; Score 16; DB 2; Length 294;
Best Local Similarity 20.0%; Pred. No. 7.8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXC 13
Db 43 CSSSSSNVC 52

RESULT 34

C71498
probable Yop translocation R - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C/Species: Chlamydia trachomatis
C/Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C/Accession: C71498
R/Stephens, R.S.; Kallman, S.; Lammell, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, S.; Science 282, 754-759, 1998
A/Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
A/Reference number: A71570; MUID:99000809; PMID:9784136
A/Accession: C71498
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-306 <ARN>
A/Cross-references: UNIPROT:084566; UNIPARC:UPI000003360; GB:AE001327; GB:AE001273; NID:94056491; P/Genetics:
A/Experimental source: serotype D, strain UW-3/Cx
A/Genes: yecR

Query Match 50.0%; Score 16; DB 2; Length 306;
Best Local Similarity 20.0%; Pred. No. 7.9;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 24 CSNACGASGC 33

RESULT 35
T00500
probable elicitor response element-binding protein WRKY3 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 31-Dec-2004
C/Accession: T00500; C64623
R/Rounsley, S.D.; Lin, X.; Kechum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Neuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.; Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: T00500
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-317 <R0U>
A/Cross-references: UNIPROT:Q22176; UNIPARC:UPI00001387D; EMBL:AC002391; NID:92642427;
A/Experimental source: cultivar Columbia
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Neuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.; Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: C84623
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-317 <STO>
A/Cross-references: UNIPARC:UPI00001387D; GB:AE002093; NID:92642432; PIDN:AAB87100.1; G/Genetics:
A/Genes: WRKY3; ATSP:T00D16.5; At2g23320
A/Map position: 2
A/Intons: 218/2; 260/2
C/Superfamily: DNA-binding protein WRKY3

Query Match 50.0%; Score 16; DB 2; Length 317;
Best Local Similarity 20.0%; Pred. No. 8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 203 CASASSGRC 212

RESULT 36
A84792
hypothetical protein At2g37380 [Imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: A84792
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Neuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.; Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: A84792
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-321 <STO>
A/Cross-references: UNIPROT:Q9ZUS8; UNIPARC:UPI00000A2519; GB:AE002093; NID:94056491; P/Genetics:
A/Genes: At2g37380
A/Map position: 2
C/Superfamily: Arabidopsis thaliana hypothetical protein At2g39370

Query Match 50.0%; Score 16; DB 2; Length 321;
Best Local Similarity 20.0%; Pred. No. 8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 306 CSRTSVSTC 315

RESULT 37
J80202
paired-box-containing Pax4 protein - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Jul-2004
C/Accession: J80202
R/Tokuyama, Y.; Yaoui, K.; Sakurai, K.; Hashimoto, N.; Saito, Y.; Kanatsuka, A.; Biochem. Biophys. Res. Commun. 248, 153-156, 1998
A/Title: Molecular cloning of rat Pax4: identification of four isoforms in rat insulinoma
A/Reference number: J80202; MUID:98340866; PMID:9675102
A/Accession: J80202
A/Molecule type: mRNA
A/Residues: 1-349 <TOX>
A/Cross-references: UNIPROT:088436; UNIPARC:UPI00001313B; GB:AF053100; NID:93192973; P/Genetics:
C/Superfamily: paired box transcription factor Pax-4; homeobox homology; paired box homology
C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
F/5-129/Domain: paired box homology <PBH>
F/171-227/Domain: homeobox homology <HOX>

Query Match 50.0%; Score 16; DB 2; Length 349;
Best Local Similarity 20.0%; Pred. No. 8.3;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 290 CSSDTTSGAC 299

RESULT 38
T04007
hypothetical protein T5L19.160 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C/Accession: T04007
R/Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X. submitted to the Protein Sequence Database, March 1999
A/Reference number: Z15184
A/Accession: T04007
A/Molecule type: DNA
A/Residues: 1-380 <BEV>
A/Cross-references: UNIPROT:Q9T0G1; UNIPARC:UPI00000A0088; EMBL:AL049481
A/Experimental source: cultivar Columbia; BAC clone T5L19
C/Genetics:
A/Map position: 4
A/Intons: 38/3; 83/3; 92/3; 123/1; 137/3; 170/3; 190/1; 207/3; 241/3; 263/3; 301/2; 358/1
A/Note: T5L19.160

Query Match 50.0%; Score 16; DB 2; Length 380;
Best Local Similarity 20.0%; Pred. No. 8.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 6 CSTASSSGC 15

RESULT 39
T26077
hypochemical protein W02A2.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T26077
R:Ainscough, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z20148
A:Accession: T26077
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-397 <WIL>
A:Cross-references: UNIPROT:Q9XUB6; UNIPARC:UPI0000076467; EMBL:Z82286; PIDD: CAB05306.1;
C:Genetics:
A:Gene: CESP:W02A2.3
A:Map position: 4
A:Introns: 15/1; 67/2; 101/3; 259/1; 383/3
C:Superfamily: gliadin

Query Match 50.0%; Score 16; DB 2; Length 397;
Best Local Similarity 20.0%; Pred. No. 8.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 10 CATLAALASC 19

RESULT 40
S41514
RAD52 protein homolog - chicken
C:Species: Gallus gallus (chicken)
C:Date: 25-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 31-Dec-2004
C:Accession: S41514
R:Bezubova, O.Y.; Schmidt, H.; Ostermann, K.; Heyer, W.D.; Buerstedt, J.M.
Nucleic Acids Res. 21, 5945-5949, 1993
A:Title: Identification of a chicken RAD52 homologue suggests conservation of the RAD52
A:Reference number: S41514; MUID:94119695; PMID:8290357
A:Accession: S41514
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-422 <BEZ>
A:Cross-references: UNIPROT:p39022; UNIPARC:UPI0000033034
C:Superfamily: human nuclear cap-binding protein

Query Match 50.0%; Score 16; DB 2; Length 422;
Best Local Similarity 20.0%; Pred. No. 8.8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 17 CTSTNSVAC 26

RESULT 41
C81039
lipopolysaccharide biosynthesis protein, probable MHA1818 [imported] - Neisseria meningi
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: C81039
R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Bisen, J.A.

Hickey, B.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
zi, H.; Qin, H.; Vamathavan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grand, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: C81039
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-473 <TET>
A:Cross-references: UNIPROT:Q9JXZ7; UNIPARC:UPI0000030867; GB:AE002531; GB:AE002098; NID
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1818

Query Match 50.0%; Score 16; DB 2; Length 473;
Best Local Similarity 20.0%; Pred. No. 9.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 421 CTTSSAAYTC 430

RESULT 42
C81984
probable lipopolysaccharide biosynthesis translocase NMA0643 [imported] - *Neisseria men*
C:Species: *Neisseria meningitidis*
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: C81984
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: C81984
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-473 <PAR>
A:Cross-references: UNIPROT:Q9JTW9; UNIPARC:UPI0000030868; GB:AL162753; GB:AL157959; NID
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA0641; NMA0643

Query Match 50.0%; Score 16; DB 2; Length 473;
Best Local Similarity 20.0%; Pred. No. 9.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 421 CLASSAAYTC 430

RESULT 43
T36342
probable glutamate decarboxylase - *Streptomyces coelicolor*
C:Species: *Streptomyces coelicolor*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36342
R:Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z21575
A:Accession: T36342
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-475 <OLI>
A:Cross-references: UNIPROT:Q9XG45; UNIPARC:UPI0000030808; EMBL:AL049841; PIDD: CAB42769.
C:Genetics:
A:Gene: gad; SCOEDB:SCB9.23
C:Superfamily: *Escherichia coli* glutamate decarboxylase

Query Match 50.0%; Score 16; DB 2; Length 475;

Best Local Similarity 20.0%; Pred. No. 9.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 131 CSTGSSSAC 140

RESULT 44

B66460 hypothetical protein F14M2.19 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: B66460

R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luro, J.S.; Malt, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: B66460

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-486 <STO>

A:Cross-references: UNIPROT:Q9LQ21; UNIPARC:UPI00000A7F31; GB:AE005172; NID:G9665100; PI

C:Genetics:

A:Map position: 1

Query Match 50.0%; Score 16; DB 2; Length 486;

Best Local Similarity 20.0%; Pred. No. 9.3;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 48 CUSSTGATTC 57

RESULT 45

A55207 glycerol-3-phosphate dehydrogenase glpd (EC 1.1.1.-) - Pseudomonas aeruginosa

C:Species: Pseudomonas aeruginosa

C:Date: 05-May-1995 #sequence_revision 05-May-1995 #text_change 09-Jul-2004

C:Accession: A55207

R:Schweizer, H.P.; Po, C.

J. Bacteriol. 176, 2184-2193, 1994

A:Title: Cloning and nucleotide sequence of the glpd gene encoding sn-glycerol-3-phospha

A:Reference number: A55207; MUID:94209216; PMID:8157588

A:Accession: A55207

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-510 <SCH>

A:Cross-references: UNIPROT:P52111; UNIPARC:UPI000016FCAD; GB:L06231; NID:G450377; PIDN:

C:Genetics:

A:Gene: glpd

C:Superfamily: glycerol-3-phosphate dehydrogenase (aerobic)

C:Keywords: oxidoreductase

Query Match 50.0%; Score 16; DB 2; Length 510;

Best Local Similarity 20.0%; Pred. No. 9.4;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 194 CTVSAAATAC 203

RESULT 46

G31277

guinate transport protein - Neurospora crassa (tentative sequence)

N:Alternate names: guinate transporter

C:Species: Neurospora crassa

C:Date: 26-Apr-1989 #sequence_revision 26-Apr-1989 #text_change 09-Jul-2004

C:Accession: S04254; G31277

R:Geever, R.F.; Hulet, L.; Baum, J.A.; Tyler, B.M.; Patel, V.B.; Rutledge, B.J.; Case, P.

J. Mol. Biol. 207, 15-34, 1989

A:Title: DNA sequence, organization and regulation of the ga gene cluster of Neurospora

A:Reference number: S04250; MUID:89293848; PMID:2525625

A:Accession: S04254

A:Molecule type: DNA

A:Residues: 1-537 <GE2>

A:Cross-references: UNIPROT:P11636; UNIPARC:UPI0000132E4B; EMBL:X14603; NID:G3060; PIDN:

C:Genetics:

A:Gene: ga-Y

C:Superfamily: maltose transport protein MAL61

C:Keywords: transmembrane protein

F:22-42/Domain: transmembrane #status predicted <TM01>

F:67-87/Domain: transmembrane #status predicted <TM02>

F:99-119/Domain: transmembrane #status predicted <TM03>

F:132-155/Domain: transmembrane #status predicted <TM04>

F:161-181/Domain: transmembrane #status predicted <TM05>

F:195-215/Domain: transmembrane #status predicted <TM06>

F:286-306/Domain: transmembrane #status predicted <TM07>

F:324-344/Domain: transmembrane #status predicted <TM08>

F:356-376/Domain: transmembrane #status predicted <TM09>

F:390-410/Domain: transmembrane #status predicted <TM10>

F:427-447/Domain: transmembrane #status predicted <TM11>

F:459-479/Domain: transmembrane #status predicted <TM12>

Query Match 50.0%; Score 16; DB 2; Length 537;

Best Local Similarity 20.0%; Pred. No. 9.6;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 24 CAIASFASC 33

RESULT 47

B38418 jockey protein 1 - fruit fly (Drosophila funebris)

C:Species: Drosophila funebris

C:Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 21-Feb-1997

C:Accession: B38418

R:Witzgib, L.U.; Mazo, A.M.

Proc. Natl. Acad. Sci. U.S.A. 87, 9216-9220, 1990

A:Title: Evidence for horizontal transmission of the mobile element jockey between dipta

A:Reference number: A38418; MUID:91067678; PMID:1701254

A:Accession: B38418

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-580 <MT2>

A:Cross-references: UNIPARC:UPI000017BB6C; GB:M38437

C:Genetics:

A:Gene: FlyBase:Dfunt/jockey

A:Cross-references: FlyBase:FBgn0012310

Query Match 50.0%; Score 16; DB 2; Length 580;

Best Local Similarity 20.0%; Pred. No. 9.9;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 78 CTVTTSQASC 87

RESULT 48

T22518 hypothetical protein FS2H3.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T22518

R.Gardner, A.
submitted to the EMBL Data Library, October 1995
A:Reference number: Z19575
A:Accession: T22518
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-638 <WIL>
A:Cross-references: UNIPROT:Q20660; UNIPARC:UPI000013C10D; EMBL:Z66512; PIDD:CAA91322.1;
A:Experimental source: clone F52H3
C:Genetics:
A:Gene: CESP:F52H3.2
A:Map position: 2
A:Introns: 67/1; 157/2; 285/2; 420/1; 566/3
C:Superfamily: gida protein

Query Match 50.0%; Score 16; DB 2; Length 638;
Best Local Similarity 20.0%; Pred. No. 10;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXC 13
DB 29 CESAARARC 38

RESULT 49
C81345
Hypothetical protein Cj0736 [imported] - Campylobacter jejuni (strain NCTC 11168)
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: C81345
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillif, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell, Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: C81345
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-642 <PAR>
A:Cross-references: UNIPROT:Q9PPH1; UNIPARC:UPI00000C1D12; GB:AL139076; GB:AL111168; NID
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj0736

Query Match 50.0%; Score 16; DB 2; Length 642;
Best Local Similarity 20.0%; Pred. No. 10;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXC 13
DB 375 CSSAAVSSC 384

RESULT 50
T37581
Probable serine-threonine-protein kinase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T37581
R:Badcock, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z21726
A:Accession: T37581
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-648 <BAD>
A:Cross-references: UNIPARC:UPI000017B219; EMBL:Z70721; PIDD:CAA94704.1; GSPDB:GN00066;
C:Genetics:
A:Gene: SPDB:SPAC12B10.14C
A:Map position: 1

Query Match 50.0%; Score 16; DB 2; Length 648;

Best Local Similarity 20.0%; Pred. No. 10;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 4 CXXXXXXXXXC 13
DB 21 COSNASSAC 30

Search completed: January 4, 2006, 16:10:17
Job time : 25.5565 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 ; Search time 104.487 Seconds

(without alignments)
108.037 Million cell updates/sec

Title: US-09-932-322-4

Perfect score: 32
Sequence: 1 XXXXXXXXXXXXXXX 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 1000 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	56.2	213	2	Q96FM3 HUMAN
2	18	56.2	263	2	Q586F4 TRYRYP
3	18	56.2	269	2	Q7PR07 ANOGA
4	18	56.2	281	2	Q5TX10 ANOGA
5	18	56.2	313	1	YMS8 YEAST
6	18	56.2	357	2	Q6ZDR4 ORYSA
7	18	56.2	362	2	Q7QJ10 ANOGA
8	18	56.2	480	2	Q9RV87 DEIRA
9	18	56.2	653	1	TMCC1 HUMAN
10	18	56.2	1005	2	Q5H317 XANOR
11	18	56.2	1061	2	Q4QCX7 LEIMA
12	17	53.1	52	2	Q8LDP6 HORVU
13	17	53.1	122	2	Q5DCN3 SCHJA
14	17	53.1	128	2	Q96LJ4 HUMAN
15	17	53.1	139	2	Q6ZKM7 ORYSA
16	17	53.1	155	2	Q6H482 ORYSA
17	17	53.1	160	2	Q657M0 ORYSA
18	17	53.1	160	2	Q4T4H1 TETNG
19	17	53.1	219	2	Q825E8 STRAM
20	17	53.1	220	2	Q66152 GVRU
21	17	53.1	256	2	Q6YTS2 ORYSA
22	17	53.1	260	2	Q850Z3 ORYSA
23	17	53.1	304	2	Q5TVE4 ANOGA
24	17	53.1	344	2	Q582B8 TRYRYP
25	17	53.1	361	2	Q57220 TRYRYP
26	17	53.1	364	2	Q4YW87 PLABE
27	17	53.1	368	2	Q70RD3 GEBHY
28	17	53.1	376	1	CBID AGRTS
29	17	53.1	388	2	Q7PY13 ANOGA
30	17	53.1	389	2	Q21081 CAENORHABDI
31	17	53.1	411	2	Q67U23 ORYSA

32	17	53.1	437	2	Q414Q8 GIBBE
33	17	53.1	471	2	Q9VMG7 DROME
34	17	53.1	475	2	Q7X7A4 ORYSA
35	17	53.1	491	2	Q4P4C7 USHMA
36	17	53.1	520	2	Q4XZV5 PLABE
37	17	53.1	556	2	Q803D5 BRABE
38	17	53.1	568	2	Q5GXM1 XANOR
39	17	53.1	626	2	Q4N2Y7 THEPA
40	17	53.1	630	2	Q69YV8 ORYSA
41	17	53.1	634	1	HSP71 LEIMA
42	17	53.1	635	2	Q4Q747 LEIMA
43	17	53.1	660	2	Q4Q740 LEIMA
44	17	53.1	662	2	Q4Q744 LEIMA
45	17	53.1	662	2	Q4Q745 LEIMA
46	17	53.1	691	2	Q60E18 ORYSA
47	17	53.1	699	2	Q819P8 TRYRYP
48	17	53.1	705	2	Q516X2 CHLAB
49	17	53.1	734	2	Q5NKE3 DROME
50	17	53.1	784	2	Q4YNR2 PLABE
51	17	53.1	843	2	Q80GV2 HPEVO
52	17	53.1	861	2	Q55NL5 CRYNE
53	17	53.1	943	1	BLI4 CABEL
54	17	53.1	1089	2	Q4Q9C2 LEIMA
55	17	53.1	1103	2	Q4PDJ1 USHMA
56	17	53.1	1200	2	Q4RFR1 TETNG
57	17	53.1	1347	2	Q6CC16 YARLI
58	17	53.1	1749	2	Q4V217 BURMA
59	17	53.1	1835	2	Q4S914 TETNG
60	17	53.1	2704	1	G168 PARER
61	17	53.1	2717	1	Q94710 PARTE
62	17	53.1	2721	2	Q76973 PARTE
63	17	53.1	2729	2	Q6P0K6 PARTE
64	16	50.0	68	2	Q6JL92 NEIHO
65	16	50.0	71	2	Q9HPK1 HALBA
66	16	50.0	80	2	Q6IGD0 DROME
67	16	50.0	85	2	Q07207 MYCBO
68	16	50.0	85	2	Q7RY24 MYCBO
69	16	50.0	85	2	Q7NSB4 CHRVO
70	16	50.0	88	2	Q5Z713 ORYSA
71	16	50.0	98	2	Q5Z2E5 NOCHA
72	16	50.0	103	2	Q5YF63 TRYRYP
73	16	50.0	104	2	Q39084 ARAYH
74	16	50.0	105	2	Q57986 PYRHO
75	16	50.0	108	2	Q9YBBS ARABE
76	16	50.0	108	2	Q85429 YCALUD
77	16	50.0	110	2	Q6Y9M8 LIGOV
78	16	50.0	112	2	Q7PUV4 ANOGA
79	16	50.0	112	2	Q95M57 BOVIN
80	16	50.0	112	2	Q8V6M5 YVIRU
81	16	50.0	120	2	Q6ID11 ARAYH
82	16	50.0	123	2	Q64WT9 ARAYH
83	16	50.0	125	2	Q59Z22 CANAL
84	16	50.0	126	2	Q7M325 PIGI
85	16	50.0	126	2	Q6YUC3 ORYSA
86	16	50.0	126	2	Q8CLV7 YERPE
87	16	50.0	129	2	Q7RUV2 NEURC
88	16	50.0	133	2	Q6UNK8 BOMBO
89	16	50.0	133	2	Q8C7T1 ARAYH
90	16	50.0	137	2	Q9HAI1 HUMAN
91	16	50.0	137	2	Q24987 GIALA
92	16	50.0	137	2	Q8WLI9 DROME
93	16	50.0	142	2	Q08497 HUMAN
94	16	50.0	147	1	IRI3 HGMYA
95	16	50.0	147	2	Q6RE18 SNCOA
96	16	50.0	149	2	Q4RZ65 TETNG
97	16	50.0	149	2	Q4S237 TETNG
98	16	50.0	151	2	Q5H3T3 XANOR
99	16	50.0	153	1	GAMB BSNV
100	16	50.0	157	2	Q8LI53 ORYSA
101	16	50.0	158	2	Q8CC93 MOUSE
102	16	50.0	161	2	Q23659 CABEL
103	16	50.0	162	2	Q91G16 ADEP3
104	16	50.0	164	2	Q7LZW7 CVB

Q414Q8 GIBBERELLA	Q414Q8 GIBBERELLA
Q9VMG7 DROSOPHILA	Q9VMG7 DROSOPHILA
Q7X7A4 ORYSA	Q7X7A4 ORYSA
Q4P4C7 USHLAGO MA	Q4P4C7 USHLAGO MA
Q4XZV5 PLABODIUM	Q4XZV5 PLABODIUM
Q803D5 BRACHYDANIO	Q803D5 BRACHYDANIO
Q5GXM1 XANTHOMONAS	Q5GXM1 XANTHOMONAS
Q4N2Y7 THELIERIA P	Q4N2Y7 THELIERIA P
Q69YV8 ORYSA	Q69YV8 ORYSA
HSP71 LEISHMANIA	HSP71 LEISHMANIA
Q4Q747 LEISHMANIA	Q4Q747 LEISHMANIA
Q4Q740 LEISHMANIA	Q4Q740 LEISHMANIA
Q4Q744 LEISHMANIA	Q4Q744 LEISHMANIA
Q4Q745 LEISHMANIA	Q4Q745 LEISHMANIA
Q60E18 ORYSA	Q60E18 ORYSA
Q819P8 TRYRYP	Q819P8 TRYRYP
Q516X2 CHLAMYDOPHIL	Q516X2 CHLAMYDOPHIL
Q5NKE3 DROSOPHILA	Q5NKE3 DROSOPHILA
Q4YNR2 PLABODIUM	Q4YNR2 PLABODIUM
Q80GV2 HEPATITIS B	Q80GV2 HEPATITIS B
Q55NL5 CRYPTOCOCCUS	Q55NL5 CRYPTOCOCCUS
P51559 CAENORHABDI	P51559 CAENORHABDI
Q4Q9C2 LEISHMANIA	Q4Q9C2 LEISHMANIA
Q4PDJ1 USHLAGO MA	Q4PDJ1 USHLAGO MA
Q4RFR1 TETRADON N	Q4RFR1 TETRADON N
Q6CC16 YARROWIA I	Q6CC16 YARROWIA I
Q4V217 BURKHOLDERIA	Q4V217 BURKHOLDERIA
Q4S914 TETRADON N	Q4S914 TETRADON N
P17053 PARAMECIUM	P17053 PARAMECIUM
Q94710 PARAMECIUM	Q94710 PARAMECIUM
Q76973 PARAMECIUM	Q76973 PARAMECIUM
Q6P0K6 PARAMECIUM	Q6P0K6 PARAMECIUM
Q6JL92 NEISSERIA G	Q6JL92 NEISSERIA G
Q9HPK1 HALOBACTERI	Q9HPK1 HALOBACTERI
Q6IGD0 DROSOPHILA	Q6IGD0 DROSOPHILA
Q07207 MYCOBACTERI	Q07207 MYCOBACTERI
Q7RY24 MYCOBACTERI	Q7RY24 MYCOBACTERI
Q7NSB4 CHROMOBACTE	Q7NSB4 CHROMOBACTE
Q5Z713 ORYSA	Q5Z713 ORYSA
Q5Z2E5 NOCARDIA FA	Q5Z2E5 NOCARDIA FA
Q5YF63 ROCK DREAM	Q5YF63 ROCK DREAM
Q39084 ARABIDOPSIS	Q39084 ARABIDOPSIS
Q57986 PYROCOCCUS	Q57986 PYROCOCCUS
Q9YBBS ACETOXYLUM P	Q9YBBS ACETOXYLUM P
Q85429 MYCOBACTERI	Q85429 MYCOBACTERI
Q6Y9M8 LIGUSTRUM O	Q6Y9M8 LIGUSTRUM O
Q7PUV4 ANOPHELES G	Q7PUV4 ANOPHELES G
Q95M57 BOE TAURUS	Q95M57 BOE TAURUS
Q8V6M5 HALOVLINUS H	Q8V6M5 HALOVLINUS H
Q6ID11 ARABIDOPSIS	Q6ID11 ARABIDOPSIS
Q64WT9 ARABIDOPSIS	Q64WT9 ARABIDOPSIS
Q59Z22 CANDIDA ALB	Q59Z22 CANDIDA ALB
Q7M325 SUS SCROFA	Q7M325 SUS SCROFA
Q6YUC3 ORYSA SATIVA	Q6YUC3 ORYSA SATIVA
Q8CLV7 YERSINIA PE	Q8CLV7 YERSINIA PE
Q7RUV2 NEUROSPORA	Q7RUV2 NEUROSPORA
Q6UNK8 BOMBYX MORI	Q6UNK8 BOMBYX MORI
Q8C7T1 ARABIDOPSIS	Q8C7T1 ARABIDOPSIS
Q9HAI1 HUMO SAPIEN	Q9HAI1 HUMO SAPIEN
Q24987 GLIADINIA IAM	Q24987 GLIADINIA IAM
Q8WLI9 DIOSPHILIA	Q8WLI9 DIOSPHILIA
Q08497 HUMO SAPIEN	Q08497 HUMO SAPIEN
P16611 HUMO CYTOM	P16611 HUMO CYTOM
Q6RE18 RHODOCOCCUS	Q6RE18 RHODOCOCCUS
Q4RZ65 TETRADON N	Q4RZ65 TETRADON N
Q4S237 TETRADON N	Q4S237 TETRADON N
Q5H3T3 XANTHOMONAS	Q5H3T3 XANTHOMONAS
Q80874 BARLEY STRI	Q80874 BARLEY STRI
Q8LI53 ORYSA SATIVA	Q8LI53 ORYSA SATIVA
Q8CC93 MUS MUSCULU	Q8CC93 MUS MUSCULU
Q23659 CAENORHABDI	Q23659 CAENORHABDI
Q91G16 PORCINE ADE	Q91G16 PORCINE ADE
Q7LZW7 CHIRYSAANTHEM	Q7LZW7 CHIRYSAANTHEM

397	16	50.0	852	2	Q9HE44_NEUCR	Q9HE44_neutropora	470	16	50.0	2128	2	Q5ONG8_EWYH1	Q5ONG8_entamoeba h
398	16	50.0	852	2	Q70010_9H1V1	Q70010 human immun	471	16	50.0	2233	2	Q94711_PAR1B	Q94711_parametium
399	16	50.0	854	2	Q5EEC9_9H1V1	Q5EEC9 human immun	472	16	50.0	2273	2	Q5ICB9_MACFA	Q5ICB9_macaca fasc
400	16	50.0	858	2	Q8AE09_9H1V1	Q8AE09 human immun	473	16	50.0	2377	2	Q7W080_BORBR	Q7W080_bordeella
401	16	50.0	859	2	Q4QG27_LE1MA	Q4QG27 leishmania	474	16	50.0	2397	2	Q27167_PARTE	Q27167_parametium
402	16	50.0	859	2	Q8ADX7_9H1V1	Q8ADX7 human immun	475	16	50.0	2397	2	Q67Y44_PARPR	Q67Y44_parametium
403	16	50.0	859	2	Q8UMF8_9H1V1	Q8UMF8 human immun	476	16	50.0	2533	2	P90569_PARTE	P90569_parametium
404	16	50.0	860	2	Q998E5_9H1V1	Q998E5 human immun	477	16	50.0	2533	2	Q27183_PARTE	Q27183_parametium
405	16	50.0	864	2	Q4U536_9H1V1	Q4U536 human immun	478	16	50.0	2543	2	P90649_PARPR	P90649_parametium
406	16	50.0	865	2	Q73342_9H1V1	Q73342 human immun	479	16	50.0	2915	1	G156_PARPR	G156_glibberella
407	16	50.0	865	2	Q73353_9H1V1	Q73353 human immun	480	16	50.0	2986	1	Q4HMU4_GIBZE	Q4HMU4_glibberella
408	16	50.0	870	2	Q5EED7_9H1V1	Q5EED7 human immun	481	16	50.0	3596	1	RPOA_SHPV	RPOA_s HPV
409	16	50.0	877	2	Q5VG71_9H1V1	Q5VG71 human immun	482	16	50.0	3869	2	Q5CRD0_CRYPV	Q5CRD0_cryptospori
410	16	50.0	879	2	Q4W9A0_ASPFU	Q4W9A0 aspergillus	483	16	50.0	3869	2	Q86PQ3_CRYPV	Q86PQ3_cryptospori
411	16	50.0	879	2	Q4PGC4_USTMA	Q4PGC4 usciello ma	484	16	50.0	4382	2	Q5D1C6_CVH9A	Q5D1C6_sars corona
412	16	50.0	881	2	Q4N476_THRPA	Q4N476 theileria p	485	16	50.0	4382	2	Q5Y188_CVH9A	Q5Y188_sars corona
413	16	50.0	884	2	Q8Q718_9H1V1	Q8Q718 human immun	486	16	50.0	4382	2	Q692E5_CVH9A	Q692E5_sars corona
414	16	50.0	889	2	Q81YQ3_PLAF7	Q81YQ3 plasmodium	487	16	50.0	4382	2	Q6JH39_CVH9A	Q6JH39_sars corona
415	16	50.0	897	2	Q5ESQ4_CRYNE	Q5ESQ4 cryptococcu	488	16	50.0	4382	2	Q6JH47_CVH9A	Q6JH47_sars corona
416	16	50.0	897	2	Q5KGF6_CRYNE	Q5KGF6 cryptococcu	489	16	50.0	4382	2	Q6R747_CVH9A	Q6R747_sars corona
417	16	50.0	907	2	Q9FW49_ARATH	Q9FW49 arabidopsis	490	16	50.0	4382	2	Q6RCW6_CVH9A	Q6RCW6_sars corona
418	16	50.0	911	2	Q8Q7U8_MOUSE	Q8Q7J8 mus musculi	491	16	50.0	4382	2	Q6RCX7_CVH9A	Q6RCX7_sars corona
419	16	50.0	912	2	Q4JHL4_FUGRU	Q4JHL4 fuga rubrip	492	16	50.0	4382	2	Q6RCY8_CVH9A	Q6RCY8_sars corona
420	16	50.0	918	2	Q6E7C9_GUBOC	Q6E7C9 oikopleura	493	16	50.0	4382	2	Q6RCZ9_CVH9A	Q6RCZ9_sars corona
421	16	50.0	920	2	Q7XGM4_ORISA	Q7XGM4 oryza sativ	494	16	50.0	4382	2	Q6RD10_CVH9A	Q6RD10_sars corona
422	16	50.0	920	2	Q94GU3_ORISA	Q94GU3 oryza sativ	495	16	50.0	4382	2	Q6RD21_CVH9A	Q6RD21_sars corona
423	16	50.0	929	2	Q5EED8_9H1V1	Q5EED8 human immun	496	16	50.0	4382	2	Q6RDJ3_CVH9A	Q6RDJ3_sars corona
424	16	50.0	934	2	Q87R90_VIBPA	Q87R90 vibrio para	497	16	50.0	4382	2	Q6RD43_CVH9A	Q6RD43_sars corona
425	16	50.0	935	2	Q4P9H9_USTMA	Q4P9H9 usciello ma	498	16	50.0	4382	2	Q6RD54_CVH9A	Q6RD54_sars corona
426	16	50.0	939	2	Q5RRH6_BRARE	Q5RRH6 brachydanio	499	16	50.0	4382	2	Q6RD65_CVH9A	Q6RD65_sars corona
427	16	50.0	967	2	Q4T490_TERNG	Q4T490 tetraodon n	500	16	50.0	4382	2	Q6VA79_CVH9A	Q6VA79_sars corona
428	16	50.0	975	2	P91357_CABEL	P91357 caenorhabdi	501	16	50.0	4382	2	Q6VA90_CVH9A	Q6VA90_sars corona
429	16	50.0	999	1	SMG_DROME	SMG2372 dirosophila	502	16	50.0	4382	2	Q6VAL1_CVH9A	Q6VAL1_sars corona
430	16	50.0	1031	2	Q65500_ARATH	Q65500 arabidopsis	503	16	50.0	4382	2	Q6WGN0_CVH9A	Q6WGN0_sars corona
431	16	50.0	1038	2	Q583N6_9TRYP	Q583N6 trypanosoma	504	16	50.0	5560	1	SPEN_DROME	SPEN_DROME
432	16	50.0	1039	2	Q4P295_USTMA	Q4P295 usciello ma	505	16	50.0	6880	2	Q6S8D8_CVH9A	Q6S8D8_sars corona
433	16	50.0	1069	2	Q592X1_CANAL	Q592X1 candida alb	506	16	50.0	7073	1	RIAB_CVH9A	RIAB_CVH9A
434	16	50.0	1069	2	Q9J3F2_CABEL	Q9J3F2 caenorhabdi	507	16	50.0	7073	2	Q5D1C7_CVH9A	Q5D1C7_sars corona
435	16	50.0	1071	2	Q61ZG1_CABEL	Q61ZG1 caenorhabdi	508	16	50.0	7073	2	Q5Y189_CVH9A	Q5Y189_sars corona
436	16	50.0	1105	2	Q20371_CABEL	Q20371 caenorhabdi	509	16	50.0	7073	2	Q692E6_CVH9A	Q692E6_sars corona
437	16	50.0	1136	2	Q92C14_LISIN	Q92C14 listeria in	510	16	50.0	7073	2	Q6JH40_CVH9A	Q6JH40_sars corona
438	16	50.0	1154	2	Q9GQ46_GIALA	Q9GQ46 giardia lam	511	16	50.0	7073	2	Q6JH48_CVH9A	Q6JH48_sars corona
439	16	50.0	1168	2	Q6AVL1_ORISA	Q6AVL1 oryza sativ	512	16	50.0	7073	2	Q6R7Y8_CVH9A	Q6R7Y8_sars corona
440	16	50.0	1191	2	Q8TOL3_DROME	Q8TOL3 dirosophila	513	16	50.0	7073	2	Q6RPP9_CVH9A	Q6RPP9_sars corona
441	16	50.0	1194	2	Q4FX33_LE1MA	Q4FX33 leishmania	514	16	50.0	7073	2	Q6UZF1_CVH9A	Q6UZF1_sars corona
442	16	50.0	1194	2	Q76M68_RAT	Q76M68 rattus norv	515	16	50.0	7073	2	Q6UZE5_CVH9A	Q6UZE5_sars corona
443	16	50.0	1227	2	Q6IFU2_SCHMA	Q6IFU2 schistosoma	516	16	50.0	7073	2	Q6V586_CVH9A	Q6V586_sars corona
444	16	50.0	1243	2	Q8QTP3_WSSV	Q8QTP3 white spot	517	16	50.0	7073	2	Q6RCW7_CVH9A	Q6RCW7_sars corona
445	16	50.0	1243	2	Q8VAP5_WSSV	Q8VAP5 white spot	518	16	50.0	7073	2	Q6RCX8_CVH9A	Q6RCX8_sars corona
446	16	50.0	1243	2	Q91LM1_WSSV	Q91LM1 white spot	519	16	50.0	7073	2	Q6RCY9_CVH9A	Q6RCY9_sars corona
447	16	50.0	1270	2	Q5UXR6_HUMAN	Q5UXR6 homo sapien	520	16	50.0	7073	2	Q6RD00_CVH9A	Q6RD00_sars corona
448	16	50.0	1310	1	GPI25_MOUSE	Q7C136 mus musculi	521	16	50.0	7073	2	Q6RD11_CVH9A	Q6RD11_sars corona
449	16	50.0	1318	2	Q4QDAS_LE1MA	Q4QDAS leishmania	522	16	50.0	7073	2	Q6RD22_CVH9A	Q6RD22_sars corona
450	16	50.0	1332	2	Q4FYPS_LE1MA	Q4FYPS leishmania	523	16	50.0	7073	2	Q6RD33_CVH9A	Q6RD33_sars corona
451	16	50.0	1353	2	Q4QBR1_LE1MA	Q4QBR1 leishmania	524	16	50.0	7073	2	Q6RD44_CVH9A	Q6RD44_sars corona
452	16	50.0	1363	2	Q4L1W7_9BPRK	Q4L1W7 burkholderi	525	16	50.0	7073	2	Q6RD55_CVH9A	Q6RD55_sars corona
453	16	50.0	1367	2	Q4QAT6_LE1MA	Q4QAT6 leishmania	526	16	50.0	7073	2	Q6RD66_CVH9A	Q6RD66_sars corona
454	16	50.0	1369	2	Q4QAT7_LE1MA	Q4QAT7 leishmania	527	16	50.0	7073	2	Q6VA80_CVH9A	Q6VA80_sars corona
455	16	50.0	1380	1	CYAA_LE1DO	Q27675 leishmania	528	16	50.0	7073	2	Q6VA91_CVH9A	Q6VA91_sars corona
456	16	50.0	1467	2	Q61ON6_CABEL	Q61ON6 caenorhabdi	529	16	50.0	7073	2	Q6VAA2_CVH9A	Q6VAA2_sars corona
457	16	50.0	1474	2	Q62504_CABEL	Q62504 caenorhabdi	530	16	50.0	7073	2	Q4UDQ5_CVH9A	Q4UDQ5_sars corona
458	16	50.0	1486	2	Q4SDM0_TERNG	Q4SDM0 tetraodon n	531	16	50.0	7073	2	Q4JDR8_CVH9A	Q4JDR8_sars corona
459	16	50.0	1497	2	Q4QHE6_LE1MA	Q4QHE6 leishmania	532	16	50.0	7073	2	Q6JH47_CVH9A	Q6JH47_sars corona
460	16	50.0	1557	2	Q5STP3_CRYNE	Q5STP3 cryptococcu	533	16	50.0	7073	2	Q6JH47_CVH9A	Q6JH47_sars corona
461	16	50.0	1558	2	Q5KIP2_CRYNE	Q5KIP2 cryptococcu	534	16	50.0	7073	2	Q6JH47_CVH9A	Q6JH47_sars corona
462	16	50.0	1568	2	Q4QOPD_LE1MA	Q4QOPD leishmania	535	16	50.0	7073	2	Q6JH47_CVH9A	Q6JH47_sars corona
463	16	50.0	1755	2	Q7WC76_BORBP	Q7WC76 bordeella	536	16	50.0	7073	2	Q6JH47_CVH9A	Q6JH47_sars corona
464	16	50.0	1769	2	Q7WC76_BORBP	Q7WC76 bordeella	537	16	50.0	7073	2	Q6JH47_CVH9A	Q6JH47_sars corona
465	16	50.0	1818	1	ENC_DROME	Q8MX1 dirosophila	538	16	50.0	7073	2	Q6JH47_CVH9A	Q6JH47_sars corona
466	16	50.0	1823	2	Q5B1H9_DROME	Q5B1H9 dirosophila	539	16	50.0	7073	2	Q6JH47_CVH9A	Q6JH47_sars corona
467	16	50.0	2071	2	Q7KV61_DROME	Q7KV61 dirosophila	540	16	50.0	7073	2	Q6JH47_CVH9A	Q6JH47_sars corona
468	16	50.0	2087	2	Q4S488_TERNG	Q4S488 tetraodon n	541	16	50.0	7073	2	Q6JH47_CVH9A	Q6JH47_sars corona
469	16	50.0	2117	2	Q4QIS6_LE1MA	Q4QIS6 leishmania	542	16	50.0	7073	2	Q6JH47_CVH9A	Q6JH47_sars corona


```
981 15 46.9 207 2 Q4V6B8 DROME
982 15 46.9 207 2 Q70016 SHIDY
983 15 46.9 207 2 Q8FCJ3 ECOL6
984 15 46.9 207 2 Q8XSN4 EC057
985 15 46.9 208 2 Q6XHV9 DROVA
986 15 46.9 208 2 Q82069 SOLTU
987 15 46.9 208 2 Q4FBJ3 EUGGL
988 15 46.9 208 2 Q9AHJ3 GLACO
989 15 46.9 209 2 Q5GML2 DICDI
990 15 46.9 210 2 Q7XKL7 ORYSA
991 15 46.9 210 2 Q93CB9 MYCPA
992 15 46.9 210 2 Q39318 9HTV1
993 15 46.9 211 1 G12 ANOGA
994 15 46.9 211 1 Q7Q5V5 ANOGA
995 15 46.9 211 2 Q50SC9 ENTHI
996 15 46.9 211 2 Q6WES3 9TILE
997 15 46.9 212 2 Q50PE5 ENTHI
998 15 46.9 212 2 Q9XV06 CABEL
999 15 46.9 212 2 Q4Q6X5 LEIMA
1000 15 46.9 212 2 Q4Q6X5 LEIMA
```

ALIGNMENTS

```
RESULT 1
Q96KM3_HUMAN PRELIMINARY; PRT; 213 AA.
```

```
AC Q96KM3;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Aberrant WW domain-containing oxidoreductase.
GN Name=WWOX;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
```

```
OC NCBI_TaxId=9606;
RN [1]
```

```
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21457335; PubMed=11572989; DOI=10.1073/pnas.191175896;
RA Palge A.J.W., Taylor K.J., Taylor C., Hillier S.G., Farrington S.,
RT Scott D., Porteous D.J., Smyth J.F., Gabra H., Watson J.E.V.;
RT "WWOX: a candidate tumor suppressor gene involved in multiple tumor
RT types.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:11417-11422(2001).
```

```
DR EMBL, AF325432; AAL05451.1; -; Genomic DNA.
DR EMBL, AF325423; AAL05451.1; JOINED; Genomic DNA.
DR EMBL, AF325424; AAL05451.1; JOINED; Genomic DNA.
DR EMBL, AF325426; AAL05451.1; JOINED; Genomic DNA.
DR EMBL, AF325433; AAL05451.1; JOINED; Genomic DNA.
DR EMBL, AF325425; AAL05451.1; JOINED; Genomic DNA.
DR HSSP; Q13526; 1PIN.
```

```
DR Ensembl; ENSG00000186153; Homo sapiens.
DR InterPro; IPR0011202; WW_Rsp5_WWP.
```

```
DR Pfam; PF00397; WW; 2.
```

```
DR SMART; SM00456; WW; 2.
```

```
DR PROSITE; PS01159; WW_DOMAIN_1; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
```

```
ID 0586F4_9TRYP PRELIMINARY; PRT; 263 AA.
AC 0586F4_9TRYP;
DT 10-MAY-2005 (TRENBLrel. 30, Created)
DT 10-MAY-2005 (TRENBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=TP927.6.4820;
OS Trypanosoma brucei.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OC NCB1_TaxId=5691;
RN [1]
```

```
RP NUCLEOTIDE SEQUENCE.
```

```
RC STRAIN=GUTat10.1;
```

```
RA Shallow E., Blandin G., Bartholomeu D., Caler E., Haas B., Hannick L.,
```

```
RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
```

```
RA Khalak H.G., Salberg S., Simpson A.J., Tallon L., Van Aken S.,
```

```
RA Wanstall D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
```

```
RA Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
```

```
RN [2]
```

```
RP NUCLEOTIDE SEQUENCE.
```

```
RC STRAIN=GUTat10.1;
```

```
RA El-Sayed N.M., Khalak H., Adams M.D.;
```

```
RA Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
```

```
RN [3]
```

```
RP NUCLEOTIDE SEQUENCE.
```

```
RC STRAIN=GUTat10.1;
```

```
RA Haas B., Blandin G., El-Sayed N.;
```

```
RA Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; AC008146; AAX80300.1; -; Genomic DNA.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```

```
DR Hypothetical protein.
```


SQ SEQUENCE 269 AA; 24228 MW; 25BBFF71FD71F1F2 CRC64;
 Query Match 56.2%; Score 18; DB 2; Length 269;
 Best Local Similarity 20.0%; Pred. No. 0.47;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 149 CTSSSSTTSC 158

RESULT 4
 05TX10 ANOGA PRELIMINARY; PRT; 281 AA.
 ID 05TX10;
 AC 01-FEB-2005 (T-EMBLrel. 29, Created)
 DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
 DE ENSANGP0000027669 (Fragment).
 GN ORFNames=ENSANG00000007023;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
 OC Anophelinae; Anopheles.
 NC NCBITaxID=180454;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PEST;
 RG The Anopheles gambiae Sequence Committee;
 RT "Anopheles gambiae re-annotation."
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PEST;
 RG The Anopheles gambiae Sequence Committee;
 RT Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAAAB01008807; EAL41760.1; -; Genomic_DNA.
 FT NON_TER 1
 FT NON_TER 281
 SQ SEQUENCE 281 AA; 25341 MW; 0462E1169FB8642B CRC64;

Query Match 56.2%; Score 18; DB 2; Length 281;
 Best Local Similarity 20.0%; Pred. No. 0.48;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 107 CTAAASSTAC 116

RESULT 5
 YMS8 YEAST STANDARD; PRT; 313 AA.
 ID YMS8 YEAST
 AC 003695;
 DT 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 10-MAY-2005 (rel. 47, Last annotation update)
 DE Hypothetical 35.0 kDa protein in PRK2-HF1 intergenic region.
 GN OrderedAccession=YMR206W; ORFNames=YMR325.07;
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycaceae;
 NC NCBITaxID=4932;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=S288c / AB972;
 RX MEDLINE=97313268; PubMed=9169872;
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
 RA Jagers K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,

RA Rice P., Skelton J., Walsh S.V., Whitehead S., Barrell B.G.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
 RT XII.";
 RL Nature 387:90-93 (1997).
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; 248755; CAA8648.1; -; Genomic_DNA.
 DR PIR; S59448; S59448.
 DR GenOnline; 142881; -;
 DR Ensembl; YMR206W; Saccharomyces cerevisiae.
 DR SGD; S000004819; YMR206W.
 KW Complete proteome; Hypothetical protein.
 FT COMPBIAS 3
 FT COMPBIAS 146 149 Poly-Gln.
 FT COMPBIAS 246 252 Poly-Ser.
 SQ SEQUENCE 313 AA; 35018 MW; 9D92BFDE982577F0 CRC64;

Query Match 56.2%; Score 18; DB 1; Length 313;
 Best Local Similarity 20.0%; Pred. No. 0.5;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
 DB 245 CSSSSSSSAC 254

RESULT 6
 06ZDR4 ORYZA PRELIMINARY; PRT; 357 AA.
 ID 06ZDR4;
 AC 06ZDR4;
 DT 05-JUL-2004 (T-EMBLrel. 27, Created)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
 DE Putative transcription factor Myb protein.
 GN Name=P0481F05.20;
 OS Oryza sativa (japonica cultivar-group);
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoideae; Oryza.
 NC NCBITaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 8, PAC
 RT clone: P0481F05."
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 DR EMBL; AP004376; BAD09322.1; -; Genomic_DNA.
 DR HSSP; Q03237; IASJ.
 DR Gramene; Q6ZDR4; -;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR InterPro; IPR01287; Homeodomain-rel.
 DR InterPro; IPR001005; Myb_DNA_bd.
 DR Pfam; PF00249; Myb_DNA-binding; 2.
 DR SMART; SM00717; SANT; 2.
 DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
 DR PROSITE; PS00334; MYB_1; 1.
 DR PROSITE; PS50090; MYB_3; 2.
 KW Nuclear protein; Repeat.
 SQ SEQUENCE 357 AA; 36754 MW; 3F4F56860D3B4E22 CRC64;

Query Match 56.2%; Score 18; DB 2; Length 357;
 Best Local Similarity 20.0%; Pred. No. 0.52;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

Db 322 CSATASASC 331

RESULT 7

Q7QJ10 ANOGA PRELIMINARY; PRT; 362 AA.

AC Q7QJ10.1

DT 01-MAR-2004 (TrEMBLrel. 26, Created)

DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)

DE ENSANGP000009383 (Fragment)

GN ORFNames=ENSANG0000007023;

OS Anopheles gambiae str. PEST.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae; Anophelinae; Anopheles.

OC NCBI_TaxID=180454;

OX [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;

RG The Anopheles gambiae Sequence Committee;

RT "Anopheles gambiae re-annotation."

RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;

RG The Anopheles gambiae Sequence Committee;

RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.

CC -! CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

DR EMBL; AAAB01008807; EAA04378.2; -; Genomic_DNA.

FT NON_TER

FT 1

SQ SEQUENCE 362 AA; 32490 MW; 384376DFA31BEC0 CRC64;

Query Match

Best Local Similarity 56.2%; Score 18; DB 2; Length 362;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 187 CTAASSTAC 196

RESULT 8

Q9RV87 DEIRA PRELIMINARY; PRT; 480 AA.

AC Q9RV87.1

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DE Hypothetical protein DR1142.

GN OrderedlocusNames=DR1142.

OS Deinococcus radiodurans.

OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales; Deinococcaceae; Deinococcus.

OC NCBI_TaxID=1299;

OX [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;

RX MEDLINE=20036966; PubMed=10567266; DOI=10.1126/science.286.5444.1571;

RA White O., Eissen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D., Dodson R.U., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L., Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M., Vamathevan J.U., Lam P., McDonald L.A., Utecherack T.R., Zalewski C., Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D., Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;

RA "Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1."

RL Science 286:1571-1577(1999).

DR EMBL; AE001963; AAF10716.1; -; Genomic_DNA.

PIR; E75433; E75433.

DR TIGR; DR1142; -.

KM Complete proteome; Hypothetical protein.

SQ SEQUENCE 480 AA; 49364 MW; 18B962AD5ACA24AE CRC64;

Query Match

Best Local Similarity 56.2%; Score 18; DB 2; Length 480;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 464 CASAAATAC 473

RESULT 9

TMCC1 HUMAN STANDARD; PRT; 653 AA.

AC 094876; Q68E06; Q8IXM8;

DT 16-OCT-2001 (Rel. 40, Created)

DT 10-MAY-2005 (Rel. 47, Last sequence update)

DE Transmembrane and coiled-coil domain protein 1.

GN Name=TMCC1; Synonyms=K1AA0779;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

OC NCBI_TaxID=9606;

OX [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RC TISSUE=skin;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stachenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uddin T.B., Tashyuk S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullen S.J., Bosak S.A., McGowan P.J., McKernan K.O., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Gardia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.B., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RL [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 15-653.

RC TISSUE=Endometrial tumor;

RG The German cDNA consortium;

RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.

RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 334-653.

RC TISSUE=Brain;

RX MEDLINE=99087487; PubMed=9872452;

RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."

RL DNA Res. 5:277-286(1998).

CC -! SIMILARITY: Belongs to the TEX28 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation CC the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

```

CC -----
DR EMBL; AB018322; BAA34499.1; -; mRNA.
DR EMBL; BC039859; AAH39859.1; -; mRNA.
DR EMBL; CR749206; CAH18064.2; -; mRNA.
DR Ensemble; ENSG00000172765; Homo sapiens.
DR HGN; HGNC:29116; TMCC1.
KW Coiled coil; Transmembrane.
FT TRANSMEM 592 612 Potential.
FT TRANSMEM 625 645 Potential.
FT COILED 228 313 Potential.
FT COILED 458 576 Potential.
SQ SEQUENCE 653 AA; 72053 MW; 18B07D17B874205 CRC64;

Query Match
Best Local Similarity 56.2%; Score 18; DB 1; Length 653;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 173 CAAAAAAAC 182

RESULT 10
Q5H317_XANOR PRELIMINARY; PRT; 1005 AA.
AC 05H317;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE Transcriptional regulator.
GN Name=acok; OrderedlocusNames=XO01230;
OS Xanthomonas oryzae (pv. oryzae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=64187;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KACC10331 / KX085;
RX PubMed=15673718; DOI=10.1093/nar/gki206;
RA Lee B.-M., Park Y.-J., Park D.-S., Kang H.-W., Kim J.-G., Song E.-S.,
RA Park I.-C., Yoon U.-H., Hahn J.-H., Koo B.-S., Lee G.-B., Kim H.,
RA Park H.-S., Yoon K.-O., Kim J.-H., Jung C.-H., Koh N.-H., Seo J.-S.,
RA Go S.-J.;
RT "The genome sequence of Xanthomonas oryzae pathovar oryzae KACC10331,
RT the bacterial blight pathogen of rice.";
RL Nucleic Acids Res. 33:577-586(2005).
DR EMBL; AE013598; AAW74484.1; -; Genomic DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0017111; F:nucleoside-triphosphatase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR000792; HTH_LuxR.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR011717; TPR_4.
DR Pfam; PF00196; Gere; 1.
DR Pfam; PF00515; TPR_1; 2.
DR Pfam; PF07721; TPR_4; 1.
DR PRINTS; PR00038; HTHLUXR.
DR ProDom; PD000307; HTH_LuxR; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00421; HTH_LuxR; 1.
DR SMART; SM00026; TPR; 3.
DR PROSITE; PS00622; HTH_LuxR_FAMILY; 1.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Complete proteome.
SQ SEQUENCE 1005 AA; 110791 MW; 662961379FD37B61 CRC64;

Query Match
Best Local Similarity 56.2%; Score 18; DB 2; Length 1005;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 4 CXXXXXXXXX 13
DB 386 CASATTAAC 395

RESULT 11
Q4QCK7_LEIMA PRELIMINARY; PRT; 1061 AA.
AC Q4QCK7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Kinesin heavy chain, putative.
GN ORFNames=ImjF20.0640;
OS Leishmania major.
OC Eukaryota; Eukaryozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005259; CAJ03791.1; -; Genomic DNA.
SQ SEQUENCE 1061 AA; 113930 MW; 2C69385D06371DCD CRC64;

Query Match
Best Local Similarity 56.2%; Score 18; DB 2; Length 1061;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 192 CAATTAATSTC 201

RESULT 12
Q8LPD6_HORVU PRELIMINARY; PRT; 52 AA.
AC Q8LPD6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative glutathione-S-transferase (Fragment).
GN Name=bar2;
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OX NCBI_TaxID=4513;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Baldwin A.;
RT "The effect of thiocarbamate herbicides and the safener dichlorimid in
RT barley and wild cats.";
RL Thesis (2001), Department of Cardiff School of Biosciences, Cardiff
RL university, Cardiff, United Kingdom.
RL [2]
RP NUCLEOTIDE SEQUENCE.
RA Baldwin A., Harwood J.L., Machray G.C., Francis D., Rogers H.J.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ419772; CAD11963.1; -; mRNA.
DR HSSP; P12653; IAXD.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002197; HTH_Pis.
DR PRINTS; PR01590; HTHPIS.
KW Transferase.
FT NON_TER 1 1
SQ SEQUENCE 52 AA; 5750 MW; 7392EB72A3C70B9B CRC64;

```

Query Match 53.1%; Score 17; DB 2; Length 52;
 Best Local Similarity 20.0%; Pred. No. 1.3;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
 DB 18 CTSASSTRSC 27

RESULT 13
 OSDCN3 SCHJA PRELIMINARY; PRT; 122 AA.
 AC QSDCN3;
 DT 10-MAY-2005 (TREMBlrel. 30, Created)
 DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
 DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
 DE Hypothetical protein.
 OS Schistosoma japonicum (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeiidae;
 OC Schistosomatidae; Schistosomatidae; Schistosoma.
 OX NCBI_TaxID=6182;
 RN NUCLEOTIDE SEQUENCE.
 RA Han Z.;
 RT "The full-length cDNA sequences of Schistosoma japonicum genes."
 RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY814691; AAW26423.1; -; mRNA.
 KM Hypothetical protein.
 SQ SEQUENCE 122 AA; 13028 MW; 960B8615986958DE CRC64;

Query Match 53.1%; Score 17; DB 2; Length 122;
 Best Local Similarity 20.0%; Pred. No. 1.8;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
 DB 76 CTSSTSTSC 85

RESULT 14
 Q96LJ4 HUMAN PRELIMINARY; PRT; 128 AA.
 ID Q96LJ4;
 AC Q96LJ4;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Hypothetical protein FLJ25437.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN NUCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RA Kawakami B., Sugiyama A., Takemoto M., Suzuki Y., Hata H.,
 RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,
 RA Irie R., Otsuki T., Sato H., Nishikawa T., Nagai K., Isogai T.,
 RA Sugano S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK058166; BAB71697.1; -; mRNA.
 DR Ensembl; ENSG00000178093; Homo sapiens.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004672; F:protein kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; Kinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
 DR ATP-binding; Kinase; Nucleotide-binding; Transferase;
 SQ SEQUENCE 128 AA; 14116 MW; D30AC8C68621B0C CRC64;

Query Match 53.1%; Score 17; DB 2; Length 128;
 Best Local Similarity 20.0%; Pred. No. 1.8;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
 DB 118 CTAISSAKTC 127

RESULT 15
 Q6ZKM7 ORYSA PRELIMINARY; PRT; 139 AA.
 AC Q6ZKM7;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE Hypothetical protein OJ1118_A06.7-1.
 GN Name=OJ1118_A06.7-1;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoideae; Oryza.
 OX NCBI_TaxID=39947;
 RN NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsunoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, BAC
 clone:OJ1118_A06.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003873; BAD08805.1; -; Genomic DNA.
 DR Gramene; Q6ZKM7;
 KM Hypothetical protein.
 SQ SEQUENCE 139 AA; 14109 MW; E281360370FB63C9 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 139;
 Best Local Similarity 20.0%; Pred. No. 1.9;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
 DB 77 CAITSSSC 86

RESULT 16
 Q6H482 ORYSA PRELIMINARY; PRT; 155 AA.
 ID Q6H482;
 AC Q6H482;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
 DE Hypothetical protein OJ1119_C03.6 (Hypothetical protein
 DE B1012G11.50).
 GN Name=OJ1119_C03.6; Synonyms=B1012G11.50;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoideae; Oryza.
 OX NCBI_TaxID=39947;
 RN NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsunoto T., Hattori M., Sasaki Y., Katayose Y.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, BAC
 clone:OJ1119_C03.";
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 RN NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsunoto T., Katayose Y.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, BAC
 clone:B1012G11.";
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005969; BAD26467.1; -; Genomic DNA.
 DR EMBL; AP005903; BAD38424.1; -; Genomic DNA.

DR Gramene; Q6H482; -;
KW Hypothetical protein.
SQ SEQUENCE 155 AA; 16233 MW; 02D83658EC7B52 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 155;
Best Local Similarity 20.0%; Pred. No. 2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 25 CDTTAAADC 34

RESULT 17

0657M0_ORYSA PRELIMINARY; PRT; 160 AA.

AC 0657M0; 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein P0468H06.13 (Hypothetical protein
P0468H05.38).
GN Name=P0468H06.13; Synonyms=P0468H05.38;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
NCBI_Taxid=39947;

NUCLEOTIDE SEQUENCE.

RA Sasaki T., Maksumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikemoto M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shinokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003075; BAD44997.1; -; Genomic DNA.
DR EMBL; AP003267; BAD87274.1; -; Genomic DNA.
KW Gramene; 0657M0; -;
KW Hypothetical protein.
SQ SEQUENCE 160 AA; 17630 MW; C3A5AED5C75DA9 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 160;
Best Local Similarity 20.0%; Pred. No. 2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 105 CCGATSTASC 114

RESULT 18

04T4H1_TETNG PRELIMINARY; PRT; 160 AA.

AC 04T4H1; 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAF9659, whole genome shotgun sequence.
GN ORFName=GSTENG00007326001.
OS Tetradodon nigrovittatus (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
NCBI_Taxid=99863;

NUCLEOTIDE SEQUENCE.

RA Uallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ouat-Costat C., Bernot A.,
RA Micaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segreus B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Catellico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Contancieu J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McGowan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollins H.,
RT "Genome duplication in the teleost fish Tetradodon nigrovittatus reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).

[12]

NUCLEOTIDE SEQUENCE.

RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.

DR EMBL; CA80100659; CAF9221.1; -; Genomic DNA.
SQ SEQUENCE 160 AA; 17094 MW; 61673E4F2AE5451B CRC64;

Query Match 53.1%; Score 17; DB 2; Length 160;
Best Local Similarity 20.0%; Pred. No. 2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 115 CTTARSSTSC 124

RESULT 19

0825E8_STRAW PRELIMINARY; PRT; 219 AA.

AC 0825E8; 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative retr-family transcriptional regulator.
GN OrderedListNames=SAV7510;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
NCBI_Taxid=33903;

[1]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RC MEDLINE=21477403; PubMed=11572948; DOI=10.1038/nbt820;
RX Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RC MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RX Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RT Nat. Biotechnol. 21:526-531(2003).
DR EMBL; BA000030; BAC75221.1; -; Genomic DNA.

```
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR01647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PROSITE; PS50977; HTH_TERR_2; 1.
KM Complete proteome.
SQ SEQUENCE 219 AA; 24066 MW; 5C59ADF39430F304 CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 219;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 151 CAEAAAAAAC 160

RESULT 20
Q66152_VVIRU PRELIMINARY; PRT; 220 AA.
AC Q66152;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Coat protein.
OS Cymbidium mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flexiviridae;
OC Potexvirus.
OX NCBI_TaxID=12178;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Korean isolate;
RA MEDLINE:95278762; PubMed:7758973; DOI=10.1016/0378-1119(95)00105-F;
RA Ryu K.H., Yoon K.B., Park W.M.;
RT "Nucleotide sequence of coat protein gene of cymbidium mosaic
RL Gene 156:303-304(1995).
DR EMBL; X81051; CAA56941.1; -; Genomic_RNA.
DR PIR; JC4082; JC4082.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000052; P1vir_coat.
DR Pfam; PF00286; Virus_P-coat; 1.
DR PRINTS; PR00232; POTCARLCOAT.
DR ProDom; PD00603; P1vir_coat; 1.
KM Capsid protein.
SQ SEQUENCE 220 AA; 23719 MW; 1F40C6E7CDCCF3B8 CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 220;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 166 CSAATLTATC 175

RESULT 21
Q6YTS2_ORYSA PRELIMINARY; PRT; 256 AA.
AC Q6YTS2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein P0419H09.29.
GN Name=P0419H09.29;
OS Oryza sativa (Japonica cultivar-group). Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Streptophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
```

```
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsunoto T., Katayose Y.;
RA Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005918; BAD10719.1; -; Genomic_DNA.
DR Gramene; O6YTS2; -.
KM Hypothetical protein.
SQ SEQUENCE 256 AA; 26591 MW; 63BFCAF3D2D40EA CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 256;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 203 CSASASARC 212

RESULT 22
Q85023_ORYSA PRELIMINARY; PRT; 260 AA.
AC Q85023;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE DnaJ-like protein (putative heat shock protein).
GN Name=OSUNBD00277B08.12; Synonym=OSUNBa0078D06.16;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Ganeberg K., Jones K.M.,
RA Overton II L.L., Tsaltrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanken S.S., Riedmiller S.B., Utehrack T.T., Feldlyum T.V.,
RA Yang Q.O., Haas B.J., Suh B.B., Peterson J.J., Quackebush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Have a continuous role in plant development probably in
CC the structural organization of compartments (By similarity).
DR EMBL; AC133778; AA039854.1; -; Genomic_DNA.
DR EMBL; AC133339; AAP46230.1; -; Genomic_DNA.
DR HSSP; P25685; 1HDJ.
DR Gramene; Q85023; -.
DR GO; GO:0031072; F:heat shock protein binding; IEA.
DR GO; GO:0051082; F:unfolded protein binding; IEA.
DR GO; GO:006457; P:protein folding; IEA.
DR GO; GO:0006986; P:response to unfolded protein; IEA.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ; 1.
DR PRINTS; PR00625; DNABPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS50076; DnaJ_2; 1.
KM Chaperone; Heat shock.
SQ SEQUENCE 260 AA; 28639 MW; DDC645DAF9BA1ED CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 260;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
```

Db 19 CAASAFSAC 28

RESULT 23

Q57VE4_ANOGA PRELIMINARY; PRT; 304 AA.

AC Q57VE4;
 DT 01-FEB-2005 (TRENBLREL. 29, Created)
 DT 01-FEB-2005 (TRENBLREL. 29, Last sequence update)
 DE 01-FEB-2005 (TRENBLREL. 29, Last annotation update)
 DE ENSANGP0000025853 (Fragment).
 GN ORFNames=ENSANGG0000024781;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
 OC Anophelinae; Anopheles.
 OC NCBI_TaxId=180454;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=PEST;
 RC The Anopheles gambiae Sequence Committee;
 RT "Anopheles gambiae re-annotation."
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=PEST;
 RC The Anopheles gambiae Sequence Committee;
 RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAAB01008839; EAL41365.1; -; Genomic_DNA.
 FT NON_TER 1 304
 SQ SEQUENCE 304 AA; 29530 MW; 1F55F0FB846B16B9 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 304;
 Best Local Similarity 20.0%; Pred. No. 2.6;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
 Db 20 CSSTSSTVC 29

RESULT 24

Q582B8_9TRYP PRELIMINARY; PRT; 344 AA.

AC Q582B8;
 DT 10-MAY-2005 (TRENBLREL. 30, Created)
 DT 10-MAY-2005 (TRENBLREL. 30, Last sequence update)
 DT 10-MAY-2005 (TRENBLREL. 30, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=TB927.5.2480;
 OS Trypanosoma brucei.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OC NCBI_TaxId=5631;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=GUTat10.1;
 RA Shedin E., Blandin G., Bartholomew D., Caler E., Haas B., Hannick L.,
 RA Shallom J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,
 RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
 RA Khailak H.G., Salzberg S., Simpson A.J., Tallon L., Van Aken S.,
 RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
 RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.

NCLEOTIDE SEQUENCE.
 RP STRAIN=GUTat10.1;
 RA El-Sayed N.M., Khailak H., Adams M.D.;
 RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
 [3]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=GUTat10.1;
 RA Haas B., Blandin G., El-Sayed N.;
 RL Submitted (Apr-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC091655; AAX80451.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 344 AA; 37545 MW; A3925C9B2A48BD53 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 344;
 Best Local Similarity 20.0%; Pred. No. 2.7;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
 Db 251 CSSSTASQSC 260

RESULT 25
 Q57Z20_9TRYP PRELIMINARY; PRT; 361 AA.

AC Q57Z20;
 DT 10-MAY-2005 (TRENBLREL. 30, Created)
 DT 10-MAY-2005 (TRENBLREL. 30, Last sequence update)
 DT 10-MAY-2005 (TRENBLREL. 30, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=TB927.5.4290;
 OS Trypanosoma brucei.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OC NCBI_TaxId=5631;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=GUTat10.1;
 RC Shedin E., Blandin G., Bartholomew D., Caler E., Haas B., Hannick L.,
 RA Shallom J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,
 RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
 RA Khailak H.G., Salzberg S., Simpson A.J., Tallon L., Van Aken S.,
 RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=GUTat10.1;
 RC Shedin E., Blandin G., Bartholomew D., Caler E., Haas B., Hannick L.,
 RA Shallom J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,
 RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
 RA Khailak H.G., Salzberg S., Simpson A.J., Tallon L., Van Aken S.,
 RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.

NCLEOTIDE SEQUENCE.
 RP STRAIN=GUTat10.1;
 RA Haas B., Blandin G., El-Sayed N.;
 RL Submitted (Apr-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC116668; AAX80579.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 361 AA; 40083 MW; 95BC9BDCB3B9895 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 361;
 Best Local Similarity 20.0%; Pred. No. 2.7;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
 Db 205 CSSSALSSSC 214

RESULT 26

Q4YW87_PLABE PRELIMINARY; PRT; 364 AA.

AC Q4YW87;
 DT 13-SEP-2005 (TRENBLREL. 31, Created)
 DT 13-SEP-2005 (TRENBLREL. 31, Last sequence update)
 DT 13-SEP-2005 (TRENBLREL. 31, Last annotation update)
 DE Hypothetical protein (Fragment).
 GN ORFNames=PB105184.00.0;
 OS Plasmodium berghei.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OC NCBI_TaxId=5821;

[1]
RN NUCLEOTIDE SEQUENCE.
RP Hall N., Kairas M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carnocci D.J., Yates J.R., Kafatos F.C.,
RA Jansz C.J., Barrell B., Turner C.M.R., Waters A.P., Stinden R.S.,
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
transcriptomic, and proteomic analyses."
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CA101002130; CAH97719.1; -; Genomic_DNA.
KM Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 364 AA; 41304 MW; 1C1711BE940AA94F CRC64;
Query Match 53.1%; Score 17; DB 2; Length 364;
Best Local Similarity 20.0%; Pred. No. 2.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXC 13
DB 332 CSTSTTASBC 341
RESULT 27
Q70RD3 GERRY PRELIMINARY; PRT; 368 AA.
ID Q70RD3;
AC 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MYB1 protein.
GN Name=Myb1;
OS Gerbera hybrid cv. 'Terra Regina'.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC asterids; campanulids; Asterales; Asteraceae; Multisidaceae;
OC Gerbera.
OC NCBI_Taxid=226891;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Petal;
RX PubMed:14605235; DOI=10.1104/pp.103.026039;
RA Elomaa P., Uimari A., Mehto M., Albert V.A., Laitinen R., Teeri T.H.;
RT "Activation of anthocyanin biosynthesis in Gerbera hybrida
(Asteraceae) suggests conserved protein-protein and protein-promoter
interactions between the anciently diverged monocots and eudicots."
RT Plant Physiol. 133:1831-1842(2003).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AJ554697; CAD87007.1; -; mRNA.
DR HSSP; 003237; IASJ.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005677; F:DNA binding; IEA.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR01005; Myb DNA bd.
DR Pfam; PF00249; Myb DNA-binding; 2.
DR SMART; SM00717; SANT; 2.
DR PROSITE; PS00037; MYB_1; UNKOWN_1.
DR PROSITE; PS00334; MYB_2; 1.
DR PROSITE; PS50090; MYB_3; 2.
KM Nuclear protein; Repeat.
SQ SEQUENCE 368 AA; 40911 MW; 7C9F8CAD3AD9421 CRC64;
Query Match 53.1%; Score 17; DB 2; Length 368;
Best Local Similarity 20.0%; Pred. No. 2.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXC 13

DB 286 CATSTSTBSC 295
RESULT 28
CBID AGRTS
ID CBID AGRTS STANDARD; PRT; 376 AA.
AC 08UBD6;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Putative cobalt-precursorin-6A synthase (deacylating) (EC 2.1.1.-).
GN Name=cbid; OrderedLocustNames=Atu2795, AGP C 5073;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Agrobacterium.
OX NCBI_Taxid=176299;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Secubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Ut., Woo L.,
RA Chen Y., Paulsen I.T., Bisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Grollio B., Goldman B.S., Cao Y., Askewazi M., Halling C., Mullin L.,
RA Hounnel K., Gordon J., Vaudin M., Tatchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughy D., Scott C., Jappes C., Markelz B.,
RA Flanagan C., Crowell C., Garson J., Lomo C., Sear C., Strub G.,
RA Cleto C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
CC -!- FUNCTION: May catalyze the methylation of C-1 in cobalt-precursorin-
CC 5 and the subsequent extrusion of acetic acid from the resulting
CC intermediate to form cobalt-precursorin-6A.
CC -!- PATHWAY: Adenosylcobalamin biosynthesis; anaerobic branch of
CC corrin ring synthesis; seventh step.
CC -!- SIMILARITY: Belongs to the cbid family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AE009227; AAL43776.1; -; Genomic_DNA.
DR EMBL; AE008192; AAK8508.1; ALT_INIT; Genomic_DNA.
DR PIR; AB2920; AB2920.
DR PIR; C97694; C97694.
DR HAMAP; MF_00787; -; 1.
DR InterPro; IPR002748; Cbid.
DR Pfam; PF01888; Cbid; 1.
DR TIGRFAMs; TIGR00312; cbid; 1.
KM Cobalamin biosynthesis; Complete proteome; Transferase.
SQ SEQUENCE 376 AA; 39088 MW; D5437CB4CA463D67 CRC64;
Query Match 53.1%; Score 17; DB 1; Length 376;
Best Local Similarity 20.0%; Pred. No. 2.8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXC 13

Db 17 CAAATKAC 26

RESULT 29

Q7PY13 ANOGA PRELIMINARY; PRT; 388 AA.

AC Q7PY13;

DT 01-MAR-2004 (TrEMBLrel. 26, Created)

DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE ENSANGP0000018378 (Fragment).

GN ORFNames=ENSANG0000018589;

OS Anopheles gambiae str. PEST.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;

OC Anophelinae; Anophelinae.

NCBI_TaxID=180454;

NP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST.

RA Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

CC -!- CAUTION: The sequence shown here is derived from an

EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

preliminary data.

DR EMBL; AAA01008987; EAA01136.1; -; Genomic_DNA.

DR InterPro; IPR000195; RabGAP_TBC.

DR Pfam; PF00566; TBC_1; RabGAP_TBC.

DR PROSITE; PS50086; TBC_RABGAP; 1.

FT NON TER 388 388

SQ SEQUENCE 388 AA; 44534 MW; A3EEC567D458CCFF CRC64;

Query Match 53.1%; Score 17; DB 2; Length 388;

Best Local Similarity 20.0%; Pred. No. 2.8;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

Db 38 CSTTAFSTC 47

RESULT 30

Q21081 CAEEL PRELIMINARY; PRT; 389 AA.

AC Q21081;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Hypothetical protein K01C8.2.

GN ORFNames=K01C8.2;

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;

OC Rhabditidae; Telodierinae; Caenorhabditis.

NCBI_TaxID=6239;

NP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=Bristol N2.

RK MEDLINE=99069613; PubMed=9851916;

RG The C. elegans sequencing consortium;

RT "Genome sequence of the nematode C. elegans: a platform for

investigating biology."

RL Science 282:2012-2018(1998).

DR EMBL; Z49068; CAA8855.1; -; Genomic_DNA.

DR PIR; T23167; T23167.

DR Ensemble; K01C8.2; Caenorhabditis elegans.

DR Wormbase; WBGene00010457; K01C8.2.

DR Wormpep; K01C8.2; CE02263.

DR InterPro; IPR006150; Worm_repeat_1.

DR SMART; SM00289; WRI; 7.

KW Complete proteome; Hypothetical protein.

RA SEQUENCE 389 AA; 40335 MW; C32B3846991D17D CRC64;

Query Match 53.1%; Score 17; DB 2; Length 389;

Best Local Similarity 20.0%; Pred. No. 2.8;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

Db 123 CSSSISTSSC 132

RESULT 31

Q67U23 ORYSA PRELIMINARY; PRT; 411 AA.

AC Q67U23;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Chloroplast thylakoidal processing peptidase-like protein.

GN Name=P048D02.23;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

NCBI_TaxID=39947;

NP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.;

RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC

clone: P048D02."

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP005393; BAD38026.1; -; Genomic_DNA.

DR Gramene; Q67U23; -.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR GO; GO:0008236; F:serine-type peptidase activity; IEA.

DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.

DR InterPro; IPR006198; Peptidase_S26A.

DR Pfam; PF00717; Peptidase_S24_1.

DR PRINTS; PR00727; LEADERPTASE.

DR TIGRPFAM; TIGR02227; sigrep_1_bact. 1.

DR PROSITE; PS00501; SPASE_1_1.

DR PROSITE; PS00761; SPASE_1_3; 1.

KW Hydrolase.

SQ SEQUENCE 411 AA; 43573 MW; EE76AD564816A373 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 411;

Best Local Similarity 20.0%; Pred. No. 2.9;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

Db 25 CSSASTAGC 34

RESULT 32

Q41408 GIBZE PRELIMINARY; PRT; 437 AA.

AC Q41408;

DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Hypothetical protein.

GN ORFNames=FG07800.1;

OS Gibberella zeae PH-1.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.

NCBI_TaxID=229533;

NP NUCLEOTIDE SEQUENCE.

RC STRAIN=PH-1;

RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,

RA Arachchi H.M., Barta N., Bastien V., Bloom T., Boguslavsky L.,

RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,

RA Choepeel Y., Collamore A., Cook A., Cooke P., Corum B., Dearrellano K.,
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
 RA Erickson J., Fero S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
 RA Gardyna S., Gierke S., Graham L., Grand-Pierre N., Hafez N.,
 RA Hagopian D., Hago B., Hall J., Horton L., Hulme W., Iliev I.,
 RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karates A.,
 RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
 RA Ma L.-Y., Mabbitt R., Maclean C., MacDonald P., Major J., Manning J.,
 RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Meneus L.,
 RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
 RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
 RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
 RA Rachupka A., Ramsamy U., Raymond C., Retta R., Rice C., Rogov P.,
 RA Roman J., Schauer S., Schupack R., Seaman S., Severy P., Smirnov S.,
 RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs W.,
 RA Talamas J., Testfaye S., Theodore J., Topham K., Travers M.,
 RA Vassiliev H., Venkataraman V.S., Vitel R., Vo A., Wang S., Wilson B.,
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
 RA Lander E.;
 RA "Fusarium graminearum genome sequence."
 RA Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AACN01000320; EAA77994.1; -; Genomic DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 437 AA; 47817 MW; 01EBA0260503394 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 437;
 Best Local Similarity 20.0%; Pred. No. 2.9;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 CXXXXXXXXX 13
 DB 199 CATLASTSSC 208

RESULT 33
 Q9VMG7_DROME PRELIMINARY; PRT; 471 AA.
 ID Q9VMG7_DROME PRELIMINARY; PRT; 471 AA.
 AC Q9VMG7;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE CG13990-PA.
 GN Name=CG13990; ORFNames=CG13990;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 OX [1]
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaratunga C.P., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Daw I., Dietz S.M.,
 RA Dodson K., Doop L.E., Domes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster A., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kratz C., Kravitz S., Kulp D., Lai Z.,
 RA Liao P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mervilov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,
 RA Palczolo M., Peltman G.S., Pan S., Pollard J., Ruit V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RT Science 287:2185-2195(2000).
 RL [2]
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=22426065; PubMed=12537568;
 RX Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Friese E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence."
 RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RL [3]
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=22426070; PubMed=12537573;
 RX Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,
 RA Patel S., Friese E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomic perspective."
 RT Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
 RL [4]
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=22426069; PubMed=12537572;
 RX Mitra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review."
 RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RL [5]
 RN NUCLEOTIDE SEQUENCE.
 RP Berkeley Drosophila Genome Project;
 RA Celniker S., Carlson J., Wan K., Pfeiffer B., Friese E., George R.,
 RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
 RA Yu C., Rubin G.;
 RT "Drosophila melanogaster release 4 sequence."
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RL [6]
 RN NUCLEOTIDE SEQUENCE.
 RP Flybase;
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 RX EMBL; AEO03612; AAF52351.2; -; Genomic DNA.
 CC Ensemble; CG13990; Drosophila melanogaster.
 DR Flybase; FBgn0040950; CG13990.
 DR GO; GO:0005576; C:extracellular region; IEA.
 DR GO; GO:0008061; F:chitin binding; IEA.
 DR GO; GO:0006030; P:chitin metabolism; IEA.
 DR InterPro; IPR002557; Chitin_bind_Pera.
 DR Pfam; PF01607; CBM_14; 2.

DR SMART: SM00494; ChtBD2; 2.
DR PROSITE: PSS0940; CHIT BIND II; 2.
SQ SEQUENCE 471 AA; 4811 MW; F316780728E9578 CRC64;
Query Match 53.1%; Score 17; DB 2; Length 471;
Best Local Similarity 20.0%; Pred. No. 3;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXX 13
DB 350 CASSTTTC 359
RESULT 34
Q7X7A4 ORYSA
ID Q7X7A4_ORYSA PRELIMINARY; PRT; 475 AA.
AC Q7X7A4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE OSJNB0060P14.14 protein (OSJNB0048E02.5 protein).
GN Name=OSJNB0060P14.14; SYNONYMS=OSJNB0048E02.5;
OC Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
NCBI_TaxID=39947;
RX NCB1_TaxID=39947;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RX PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Weng Q., Zhang L., Lu Y., Mu Y., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan D.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL663017; CAE04361.2; -; Genomic_DNA.
DR EMBL; AL606653; CAE04825.2; -; Genomic_DNA.
DR Gramene; Q7X7A4; -;
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00256; FBOX; 1.
DR PROSITE; PSS0181; FBOX; 1.
SQ SEQUENCE 475 AA; 53063 MW; 63345B3CB2D4CA5 CRC64;
Query Match 53.1%; Score 17; DB 2; Length 475;
Best Local Similarity 20.0%; Pred. No. 3;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXX 13
DB 417 CSSSTLSAC 426
RESULT 35
Q4P4C7 USTMA
ID Q4P4C7_USTMA PRELIMINARY; PRT; 491 AA.
AC Q4P4C7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UW05036.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.

OX NCB1_TaxID=237631;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521.
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adkoya E.,
RA Alt-zabra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantung P., Baldwin J., Barry A.,
RA Bayul T., Blitschey B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Bouhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheenatung Y., Chtren M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Ekins T., Engels R.,
RA Erickson J., Farina K., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gartin G., Gnerre S.,
RA Gritke A., Goyette A., Graham J., Grandbois E., Gyalten K., Hafez N.,
RA Hagopian D., Hages B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysseils M., Karlsson E.,
RA Kells C., Kieu A., Kiser P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokitsang T., Lokitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Mayor J.,
RA Manning J., Mardella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrum J., Meus L.,
RA Mestrov J., Mihalev A., Mihova T., Mikelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okafo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pignat B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan U., Sherpa N., Shi J., Smirnov S., Smith C., Sougniez C.,
RA Spencer B., Stalker U., Stange-thomann N., Stavropoulos S.,
RA Steuson K., Stone C., Stone S., Stubbs M., Talamas J., Tchinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallée D., Vassiliou H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembek L.,
RA Zimmer A., Zody M., Zander B.;
RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC DR EMBL; AACP0100186; BAK85896.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 491 AA; 51869 MW; 768A84FBB8D71BBE CRC64;
Query Match 53.1%; Score 17; DB 2; Length 491;
Best Local Similarity 20.0%; Pred. No. 3.1;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXX 13
DB 453 CCASTGSSC 462
RESULT 36
Q4XZV5 PLACH
ID Q4XZV5_PLACH PRELIMINARY; PRT; 520 AA.
AC Q4XZV5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
GN ORFNames=PC000338.02.0;
OS Plasmodium chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OC NCB1_TaxID=5825;
RN (1)

RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karris N., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutledge K., Harris B., Harris D., Churche C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carrucci D.J., Yates J.P., Kafatos F.C.,
RA Jane C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.,
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses."
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: CAJ01002036; CANT7555.1; -: Genomic_DNA.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 5.
DR PRINTS: PR00320; GPROTEINRPT.
DR SMART: SM00320; WD40; 5.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_1.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Repeat; WD repeat.
KW NON_TER 520 520
SQ SEQUENCE 520 AA; 59951 MW; FA789C11BFD99087 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 520;
Best Local Similarity 20.0%; Pred. No. 3.1;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 364 CTTSSLSSSC 373

RESULT 37
Q803D5 BRARE PRELIMINARY; PRT; 556 AA.
ID Q803D5 BRARE PRELIMINARY;
AC Q803D5
DT 01-UN-2003 (TREMBLrel. 24, Created)
DT 01-UN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Similar to adrenal secretory serine protease.,
GN ORFNames=zgc:55888;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_Taxid=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN-AB; TISSUE=whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Staciecon M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein W.J., Ushed T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman W., Madan A., Rodighiero S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP NUCLEOTIDE SEQUENCE.

RC STRAIN-AB; TISSUE=whole body;
RA Strausberg R.,
RT Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC044526; AAH44526.1; -: mRNA.
DR HSSP: P00760; IEZK.
DR Ensembl: ENSDARG00000016538; Danio rerio.
DR ZFIN: ZDB-GENE-040426-834; zgc:55888.
DR GO: GO:0004263; F-tyrosinase activity; IEA.
DR GO: GO:0008233; F-tyrosinase activity; IEA.
DR GO: GO:0004295; F-tyrosinase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR001254; Peptidase_S1_S6.
DR InterPro: IPR001314; Peptidase_S1A.
DR Pfam: PF00089; Tyrosin_2.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRYPSIN_DOM; 2.
DR PROSITE: PS00240; TRYPSIN_DOM; 2.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
KW SEQUENCE 556 AA; 61077 MW; C7B23D930547878C CRC64;

Query Match 53.1%; Score 17; DB 2; Length 556;
Best Local Similarity 20.0%; Pred. No. 3.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 493 CTHAAASTSC 502

RESULT 38
Q5GXM1_XANOR PRELIMINARY; PRT; 568 AA.
ID Q5GXM1_XANOR PRELIMINARY;
AC Q5GXM1;
DT 10-MAY-2005 (TREMBLrel. 30, Created)
DT 10-MAY-2005 (TREMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
DE Alkaline phosphatase.
GN Name=phoA; OrderedLocustNames=XCO03296;
OS Xanthomonas oryzae (pv. oryzae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_Taxid=64187;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=XACC10331 / KX085;
RX PubMed=15673718; DOI=10.1093/nar/gk1206;
RA Lee B.-M., Park Y.-J., Park D.-S., Kang H.-W., Kim J.-G., Song E.-S.,
RA Park I.-C., Yoon U.-H., Hahn J.-H., Ko B.-S., Lee G.-B., Kim H.,
RA Park H.-S., Yoon K.-O., Kim J.-H., Jung C.-H., Koh N.-H., Seo J.-S.,
RA Go S.-U.;
RT "The genome sequence of Xanthomonas oryzae pathovar oryzae XACC10331,
RT the bacterial blight pathogen of rice."
RL Nucleic Acids Res. 33:577-586(2005).
DR EMBL: AB013598; AAW76550.1; -: Genomic_DNA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR001952; ALK_phosphatase.
DR Pfam: PF00245; ALK_phosphatase; 1.
DR PRINTS: PR00113; ALKPHPTASE.
DR SMART: SM00098; alkprc; 1.
KW Complete proteome.
SQ SEQUENCE 568 AA; 60106 MW; CAE69E266ACB3A6 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 568;
Best Local Similarity 20.0%; Pred. No. 3.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 21 CASTAGSSAC 30

```

RESULT 39
ID QAN2Y7 THEPA PRELIMINARY; PRT; 626 AA.
AC QAN2Y7;
DT 13-SEP-2005 (Tremblrel. 31, Created)
DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
DE Hypothetical protein.
OC ORFNames:TP04_0205;
OS Theileria parva.
OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileridae;
OC Theileria.
OX NCBI_TaxID=5875;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Muga;
RA Gardner M.D., Bishop R., Shah T., de Villiers E.P., Carlton J.M.,
RA Hall N., Ren Q., Paulsen I.T., Pain A., Berriman M., Wilson R.J.M.,
RA Sato S., Ralph S.A., Mann D.J., Xiong Z., Shallom S.J., Weidman J.,
RA Jiang L., Lynn J., Weaver B., Shoabdi A., Maawo D., Crabtree J.,
RA Wortman J.R., Haas B., Angiuoli S.V., Creasy T.H., Lu C., Suh B.,
RA Silva J.C., Uterback T.R., Feldblyum T.V., Perrea M., Allen J.,
RA Taracha E.L.N., Salzberg S.L., White O.R., Fitzhugh H.A., Morzaria S.,
RA Venter J.C., Fraser C.M., Nene V.;
RA "Genome sequence of Theileria parva, a bovine pathogen that transforms
RT lymphocytes."
RL Science 309:134-137(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Muga;
RA Gardner M., Bishop R., Shah T., de Villiers E., Carlton J.M., Hall N.,
RA Ren Q., Paulsen I.T., Pain A., Berriman M., Wilson R.J.M., Sato S.,
RA Ralph S.A., Mann D.J., Xiong Z., Shallom S.J., Weidman J., Jiang L.,
RA Lynn J., Weaver B., Shoabdi A., Maawo D., Crabtree J., Wortman J.R.,
RA Haas B., Angiuoli S., Creasy T.H., Lu C., Suh B., Silva J.C.,
RA Uterback T., Feldblyum T., Perrea M., Allen J., Taracha E.L.,
RA Salzberg S.L., White O., Fitzhugh H.A., Morzaria S., Venter J.C.,
RA Fraser C.M., Nene V.;
RA Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAK01000004; EMBL31557.1; -; Genomic DNA.
KW Hypothetical protein.
SQ
SEQUENCE 626 AA; 72615 MW; 663085BA95D09916 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 626;
Best Local Similarity 20.0%; Pred. No. 3.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 44 CAASSSDASC 53

RESULT 40
ID Q69TV8 ORYSA PRELIMINARY; PRT; 630 AA.
AC Q69TV8;
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)
DE Receptor protein kinase-like.
GN Name=OSUNBA0029G06.34; Synonym=OSUNBA0072A21.3;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Sasaki T., Matsumoto T., Yamamoto K.;
RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC

```

```

RT clone:OSUNBA0029G06.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC Sasaki T., Matsumoto T., Yamamoto K.;
RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC
RT clone:OSUNBA0072A21.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004680; BAD35689.1; -; Genomic DNA.
DR EMBL; AP004737; BAD37734.1; -; Genomic DNA.
DR Gramene; Q69TV8; -;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR GO; GO:0004668; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR002482; LysM.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF01476; LysM; 1.
DR Pfam; PF00069; PKinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
KW kinase, Receptor.
SQ
SEQUENCE 630 AA; 67017 MW; F0EFC8DB8A63646 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 630;
Best Local Similarity 20.0%; Pred. No. 3.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 48 CSEYTTATGTC 57

RESULT 41
ID HSP71 LEIMA STANDARD; PRT; 634 AA.
AC P12076;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-FEB-2005 (Rel. 46, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Heat shock 70-related protein 1, mitochondrial precursor.
GN Name=HSP70.1;
OS Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OC NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MHOM/IL/81/Friedlin;
RA MEDLINE=93300981; PubMed=8314893;
RA Seattle S., McCrosan M.V., Smith D.F.;
RT "Expression of a mitochondrial stress protein in the protozoan
RT parasite Leishmania major."
RL J. Cell Sci. 104:1091-1100(1993).
RN [2]
RP SEQUENCE REVISION TO 461-500.
RA Smith D.F.;
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-249.
RC STRAIN=MHOM/IL/81/Friedlin;
RA MEDLINE=89345072; PubMed=2762121;
RA Seattle S., Campos A.J.R., Coulson R.M.R., Spithill T.W., Smith D.F.;
RT "A family of heat shock protein 70-related genes are expressed in the
RT promastigotes of Leishmania major."
RL Nucleic Acids Res. 17:5081-5095(1989).
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.

```

```
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X64137; CAA5498.2; -; Genomic_DNA.
DR EMBL; X14574; CAA32713.1; -; Genomic_DNA.
DR PIR; S33575; S33575.
DR HSSP; P04475; 1DKG.
DR InterPro; IPR012725; Dnak prot.
DR InterPro; IPR0101023; Hsp70.
DR PANTHER; PTHR19375; Hsp70; 1.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR ProDom; PD000089; Hsp70; 1.
DR TIGRFAMs; TIGR02350; prok_dnak; 1.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; FALSE_NEG.
DR ATP-binding; Coiled coil; Heat shock; Mitochondrion; Multigene family;
DR Nucleotide-binding; Transist peptide.
DR TRANSIT 1 20 Mitochondrion (By similarity).
DR CHAIN 21 634 Heat shock 70-related protein 1.
DR COILED 538 614 Potential.
DR SEQUENCE 634 AA; 68330 MW; 8382B1C40C108BEA CRC64;

Query Match 53.1%; Score 17; DB 1; Length 634;
Best Local Similarity 20.0%; Pred. No. 3.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
Db 7 CGSAAASAAAC 16

RESULT 42
Q04747_LEIMA PRELIMINARY; PRT; 635 AA.
ID Q04747;
AC Q04747;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Heat shock 70-related protein 1, mitochondrial, putative.
GN ORFNames=Lmjf30.2460;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Fredlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RA Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
DR EMBL; CT005267; CAJ06531.1; -; Genomic_DNA.
DR InterPro; IPR012725; Dnak prot.
DR InterPro; IPR0101023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR ProDom; PD000089; Hsp70; 1.
DR TIGRFAMs; TIGR02350; prok_dnak; 1.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; 1.
DR ATP-binding; Heat shock; Nucleotide-binding.
DR SEQUENCE 635 AA; 68948 MW; 1A11F63117CA2700 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 635;
Best Local Similarity 20.0%; Pred. No. 3.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 4 CXXXXXXXXXC 13
Db 7 CGSAAASAAAC 16

RESULT 43
Q04740_LEIMA PRELIMINARY; PRT; 652 AA.
ID Q04740;
AC Q04740;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Heat shock 70-related protein 1, mitochondrial, putative.
GN ORFNames=Lmjf30.2550;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Fredlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RA Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
DR EMBL; CT005267; CAJ06551.1; -; Genomic_DNA.
DR InterPro; IPR012725; Dnak prot.
DR InterPro; IPR0101023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR ProDom; PD000089; Hsp70; 1.
DR TIGRFAMs; TIGR02350; prok_dnak; 1.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; 1.
DR ATP-binding; Heat shock; Nucleotide-binding.
DR SEQUENCE 652 AA; 70609 MW; 52D4843JDE3C3601 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 652;
Best Local Similarity 20.0%; Pred. No. 3.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
Db 7 CGSAAASAAAC 16

RESULT 44
Q04744_LEIMA PRELIMINARY; PRT; 660 AA.
ID Q04744;
AC Q04744;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Heat shock 70-related protein 1, mitochondrial, putative.
GN ORFNames=Lmjf30.2490;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Fredlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RA Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
DR EMBL; CT005267; CAJ06539.1; -; Genomic_DNA.
DR InterPro; IPR012725; Dnak prot.
DR InterPro; IPR0101023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
```

```
DR ProDom; PD000089; Hsp70; 1.
DR TIGRFAME; TIGR02350; prok_dnak; 1.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; 1.
DR ATP-binding; Heat shock; Nucleotide-binding.
SQ SEQUENCE 660 AA; 71621 MW; 5E1D306B3D524242 CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 660;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Cy 4 CXXXXXXXXXC 13
Db 7 CGSAASAAAC 16

RESULT 45
Q40745 LEIMA PRELIMINARY; PRT; 662 AA.
ID Q40745 LEIMA PRELIMINARY; PRT; 662 AA.
AC Q40745.
DT 13-SEP-2005 (TrEMBLrel. 31, Last Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Heat shock 70-related protein 1, mitochondrial, putative.
GN ORFAMES-LmjF30.2470, LmjF30.2480;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrett B.G.;
RA Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -! SIMILARITY: Belongs to the heat shock protein 70 family.
DR EMBL; CT005267; CAJ06535.1; -; Genomic_DNA.
DR EMBL; CT005267; CAJ06535.1; -; Genomic_DNA.
DR InterPro; IPR012725; Dnak_prok.
DR InterPro; IPR001023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR ProDom; PD000089; Hsp70; 1.
DR TIGRFAME; TIGR02350; prok_dnak; 1.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; 1.
DR ATP-binding; Heat shock; Nucleotide-binding.
SQ SEQUENCE 662 AA; 71877 MW; CFC66508C46E20 CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 662;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Cy 4 CXXXXXXXXXC 13
Db 7 CGSAASAAAC 16

RESULT 46
Q60E18 ORYSA PRELIMINARY; PRT; 691 AA.
ID Q60E18 ORYSA PRELIMINARY; PRT; 691 AA.
AC Q60E18.
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative cyclic nucleotide gated ion channel.
GN Name=OSJNB0017K09.10;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
```

```
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chao Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kuo P.-I., Lee M.-C.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
OS "Oryza sativa BAC OSJNB0017K09 genomic sequence.";
RN Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC130597; AAU90233.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR000595; cAMP_binding.
DR InterPro; IPR005821; Ion_trans.
DR Pfam; PF00027; cAMP_binding; 1.
DR Pfam; PF00520; Ion_trans; 1.
DR SMART; SM00100; cAMP; 1.
DR PROSITE; PS00042; cAMP_BINDING_3; 1.
SQ SEQUENCE 691 AA; 77220 MW; D059333316F2E299 CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 691;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Cy 4 CXXXXXXXXXC 13
Db 298 CASSTAAVAC 307

RESULT 47
Q819P8 9BIVA PRELIMINARY; PRT; 639 AA.
ID Q819P8 9BIVA PRELIMINARY; PRT; 639 AA.
AC Q819P8.
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Alpha-amylase.
DR GN Name=Amv;
OS Corbicula fluminea.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
OC Corbiculoidae; Corbiculidae; Corbicula.
OX NCBI_TaxID=45949;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Da Lage J.-L., Van Wormhoudt A., Cariou M.-L.;
RN "Diversity and evolution of the alpha-amylase genes in Animals.";
RN Biologia 57:181-189(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA PubMed=14704857; DOI=10.1007/s00018-003-3334-Y;
RN "Da Lage J.-L., Feller G., Janacek S.;
RN "Horizontal gene transfer from Eukarya to bacteria and domain
RN shuffling: the alpha-amylase model.";
RN Cell. Mol. Life Sci. 61:97-109(2004).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Da Lage J.-L.;
RN Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF468016; AA017927.2; -; Genomic_DNA.
DR HSSP; P04745; LSMD.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0016798; F:hydrolase activity; acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006048; Alpha_amyl_C.
DR InterPro; IPR006049; Alpha_amyl_cat.
DR InterPro; IPR006589; Alp_amyl_cat_sub.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; Alpha-amylase; 1.
DR Pfam; PF02806; Alpha-amylase_C; 1.
```

```

DR PRINTS: PR00110; ALPHAMYLASE.
DR SMART; SM00642; Amy; 1.
DR SMART; SM00632; Amy; C; 1.
SQ SEQUENCE 699 AA; 76544 MW; 25D57008165B04CB CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 699;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 544 CTSDATSAC 553

RESULT 48
OS16X2 CHLAB
ID OS16X2 CHLAB PRELIMINARY; PRT; 705 AA.
AC OS16X2-
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Putative transport protein.
GN OrderedLocNames=CAB140;
OS Chlamydomonas abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomonada.
OK NCBI_TaxID=83555;
RN NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).
RP STRAIN=Berkley;
RC PubMed=15837807; DOI=10.1101/gr.3684805;
RX Thomson N.R., Yeates C., Bell K., Holden M.T.G., Bentley S.D.,
Livingstone M., Cerdano-Tarraga A.-M., Harris B., Doggett J.,
Ormond D., Mungall K., Clarke K., Feltwell T., Hance Z., Sanders M.,
Quail M.A., Price C., Barrett B.G., Parkhill J., Longbottom D.;
RT "The Chlamydomonas abortus genome sequence reveals an array of
RT variable proteins that contribute to interspecies variation.";
RL Genome Res. 15:629-640(2005).
DR EMBL; CR846038; CAH63598.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 705 AA; 81369 MW; 1997374B5266E2AB CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 705;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 67 CSSLSATTC 76

RESULT 49
OSNK3 DROME
ID OSNK3 DROME PRELIMINARY; PRT; 734 AA.
AC OSNK3-
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Hypothetical protein BG:DS00180.5.
GN Name=BG:DS00180.5; ORFNames=CG31766, CG31766, CG31841;
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OK NCBI_TaxID=7227;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Berkley;
RC MEDLINE=99403001; PubMed=10471707;
RX Ashburner M., Mista S., Roote J., Lewis S.E., Blazek R.G., Davis T.,
Doyle C., Galle R.F., George R.A., Harris N.L., Hartzell G.,
Harvey D.A., Hong L., Houston K.A., Hoskins R.A., Johnson G.,
Martin C., Moshrefi A.R., Palazolo M., Reese M.G., Spradling A.C.,
Tsang G., Wan K.H., Whitelaw K., Celniker S.E., Rubin G.M.,

```

```

RT "An exploration of the sequence of a 2.9-Mb region of the genome of
RT Drosophila melanogaster: the Adh region.";
RL Genetics 153:179-219(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC Celniker S.E., Agbayani A., Arcalata T.T., Baxter E., Blazek R.G.,
RA Butenoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummel S.R., Karia K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomoloni M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pacle J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA Sethi H., Smit E., Swirek R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB003408; AAF44838.1; -; Genomic_DNA.
DR FlyBase; FBgn0051765; CG31765.
DR FlyBase; FBgn0051766; CG31766.
DR FlyBase; FBgn0051841; CG31841.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004246; F:peptidyl-dipeptidase A activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR01548; Peptidase_M2.
DR Pfam; PF01401; Peptidase_M2; 1.
KW Hypothetical protein.
SQ SEQUENCE 734 AA; 79831 MW; 4C0B5993AF9FBC18 CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 734;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 681 CSSASAVSC 690

RESULT 50
Q4YNR2_PLABE
ID Q4YNR2_PLABE PRELIMINARY; PRT; 784 AA.
AC Q4YNR2-
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PB000881.03.0;
OS Plasmodium berghei.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OK NCBI_TaxID=5821;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karris M., Raine J.D., Carlton J.M., Kool J.T.W.A.,
RA Bertman M., Florens L., Jansen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Church C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Jance C.J., Barrett B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses";
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAI01003291; CAI00348.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 784 AA; 91429 MW; B76DABAD7A0C1D6 CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 784;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 363 CSSSFSFSSC 372

```


Thu Jan 5 09:14:21 2006

us-09-932-322-4.rup

Page 25

Search completed: January 4, 2006, 16:09:24
Job time : 131.487 secs

2006-01-04 16:09:24

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 ; Search time 123.183 Seconds
(without alignments)
64.204 Million cell updates/sec

Title: US-09-932-322-5
Perfect score: 34
Sequence: 1 XXXXXXXXXXXXXXXX 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21:*

- 1: Geneseqp19808:*
- 2: Geneseqp19908:*
- 3: Geneseqp20008:*
- 4: Geneseqp20018:*
- 5: Geneseqp20028:*
- 6: Geneseqp20038:*
- 7: Geneseqp20048:*
- 8: Geneseqp20058:*
- 9: Geneseqp20068:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	52.9	104	7	ABO82562 Pseudomon
2	18	52.9	110	8	ADP30723 Human sec
3	18	52.9	135	8	ADP31638 Human sec
4	18	52.9	142	7	ABO79745 Pseudomon
5	18	52.9	178	7	ABO73284 Pseudomon
6	18	52.9	264	8	ADP31527 Human sec
7	18	52.9	270	8	ADP31435 Human sec
8	18	52.9	307	4	ABE5879 Drosophila
9	18	52.9	357	8	ADP31423 Human sec
10	18	52.9	615	8	ADP31132 Human sec
11	18	52.9	627	6	ABU51580 Pseudomon
12	18	52.9	627	4	ABU48099 Pseudomon
13	18	52.9	669	8	ADP31142 Human sec
14	18	52.9	711	8	ADP31215 Human sec
15	18	52.9	739	8	ADP31196 Human sec
16	18	52.9	771	8	ADP31244 Human sec
17	18	52.9	876	8	ADP31220 Human sec
18	18	52.9	882	8	ADP31688 Human sec
19	18	52.9	925	5	ABU4246 Human sec
20	18	52.9	1044	8	ADP31517 Human sec
21	18	52.9	1086	8	ADP31175 Human sec
22	18	52.9	1113	8	ADP31508 Human sec
23	18	52.9	1128	6	ADA15725 C. elegans
24	18	52.9	1134	8	ADP30741 Human sec

25	18	52.9	1134	8	ADP30924 Human sec
26	18	52.9	1168	8	ADP31046 Human sec
27	18	52.9	1191	8	ADP30993 Human sec
28	18	52.9	1260	8	ADP31533 Human sec
29	18	52.9	1289	8	ADP30675 Human sec
30	18	52.9	1437	8	ADP31357 Human sec
31	18	52.9	1454	8	ADP31177 Human sec
32	18	52.9	1480	8	ADP30557 Human sec
33	18	52.9	1588	5	ABU09437 H. influenzae
34	18	52.9	1617	8	ADP30660 Human sec
35	18	52.9	1652	6	ADA15715 C. elegans
36	18	52.9	1725	8	ADP30654 Human sec
37	18	52.9	1833	8	ADP30642 Human sec
38	18	52.9	2088	8	ADP31178 Human sec
39	18	52.9	2127	8	ADP31327 Human sec
40	18	52.9	2484	8	ADP66590 Human sec
41	18	52.9	2508	8	ADA15721 C. elegans
42	18	52.9	2544	6	ADA15717 C. elegans
43	18	52.9	2601	6	ADA15723 C. elegans
44	18	52.9	2833	8	ADP31299 Human sec
45	18	52.9	2835	8	ADP30572 Human sec
46	18	52.9	3411	8	ADP30667 Human sec
47	18	52.9	3440	6	ABU88256 Human sec
48	18	52.9	3440	6	ABU90135 Human sec
49	18	52.9	3440	6	ABU96437 Human sec
50	18	52.9	3440	6	ABU99046 Human sec
51	18	52.9	3440	6	ABU98261 Human sec
52	18	52.9	3440	6	ABU91967 Human sec
53	18	52.9	3440	6	ABU85271 Human sec
54	18	52.9	3440	6	ABU00410 Human sec
55	18	52.9	3440	6	ABU88961 Human sec
56	18	52.9	3440	6	ABU06457 Human sec
57	18	52.9	3440	6	ABU95517 Human sec
58	18	52.9	3440	6	ABU95207 Human sec
59	18	52.9	3440	6	ABU90755 Human sec
60	18	52.9	3440	6	ABU93917 Human sec
61	18	52.9	3440	6	ABU86191 Human sec
62	18	52.9	3440	6	ABU82046 Human sec
63	18	52.9	3440	6	ABU07907 Human sec
64	18	52.9	3440	6	ABU94227 Human sec
65	18	52.9	3440	6	ABU00100 Human sec
66	18	52.9	3440	6	ABU87111 Human sec
67	18	52.9	3440	6	ABU91352 Human sec
68	18	52.9	3440	6	ABU90445 Human sec
69	18	52.9	3440	6	ABU97036 Human sec
70	18	52.9	3440	6	ABU05232 Human sec
71	18	52.9	3440	6	ABU01425 Human sec
72	18	52.9	3440	6	ADP31494 Human sec
73	18	52.9	3440	6	ADP30534 Human sec
74	18	52.9	3440	6	ADP30526 Human sec
75	18	52.9	3440	6	ABE29760 Human sec
76	18	52.9	3440	6	ADP30709 Human sec
77	18	52.9	3440	6	ABU078139 Human sec
78	18	52.9	3440	6	ABU079679 Human sec
79	18	52.9	3440	6	AAU94985 Human sec
80	18	52.9	3440	6	ADP31048 Human sec
81	18	52.9	3440	6	ADP31515 Human sec
82	18	52.9	3440	6	ADP31516 Human sec
83	18	52.9	3440	6	ADP31516 Human sec
84	18	52.9	3440	6	ADP31621 Human sec
85	18	52.9	3440	6	ADP31099 Human sec
86	18	52.9	3440	6	ADP31231 Human sec
87	18	52.9	3440	6	ADP31231 Human sec
88	18	52.9	3440	6	ADP31231 Human sec
89	18	52.9	3440	6	ADP31231 Human sec
90	18	52.9	3440	6	ADP31231 Human sec
91	18	52.9	3440	6	ADP31231 Human sec
92	18	52.9	3440	6	ADP31231 Human sec
93	18	52.9	3440	6	ADP31231 Human sec
94	18	52.9	3440	6	ADP31231 Human sec
95	18	52.9	3440	6	ADP31231 Human sec
96	18	52.9	3440	6	ADP31231 Human sec
97	18	52.9	3440	6	ADP31231 Human sec

98	17	50.0	325	7	ABM87858	Abm87858	Rice	abio	171	17	50.0	950	8	ADP31167	Adp31167	Human	sec
99	17	50.0	330	7	ADB70062	Adb70062	C. neofor		172	17	50.0	951	8	ADP30943	Adp30943	Human	sec
100	17	50.0	348	8	ADP31441	Adp31441	Human	sec	173	17	50.0	957	8	ADP31528	Adp31528	Human	sec
101	17	50.0	357	8	ADP31267	Adp31267	Human	sec	174	17	50.0	957	8	ADT07508	Adt07508	Human	col
102	17	50.0	367	8	ADP30950	Adp30950	Human	sec	175	17	50.0	960	8	ADP31471	Adp31471	Human	sec
103	17	50.0	381	8	ADP30655	Adp30655	Human	sec	176	17	50.0	960	8	ADP31470	Adp31470	Human	sec
104	17	50.0	384	8	ADP30656	Adp30656	Human	sec	177	17	50.0	981	8	ADP30547	Adp30547	Human	sec
105	17	50.0	392	8	ADP31061	Adp31061	Human	sec	178	17	50.0	996	8	ADP31538	Adp31538	Human	sec
106	17	50.0	392	8	ADP31055	Adp31055	Human	sec	179	17	50.0	1035	8	ADP31552	Adp31552	Human	sec
107	17	50.0	393	8	ADP31345	Adp31345	Human	sec	180	17	50.0	1038	8	ADP30860	Adp30860	Human	sec
108	17	50.0	402	8	ADP30742	Adp30742	Human	sec	181	17	50.0	1044	8	ADP31550	Adp31550	Human	sec
109	17	50.0	402	8	ADP31414	Adp31414	Human	sec	182	17	50.0	1048	8	ADP31642	Adp31642	Human	sec
110	17	50.0	402	8	ADP31262	Adp31262	Human	sec	183	17	50.0	1053	8	ADP30886	Adp30886	Human	sec
111	17	50.0	409	7	ABO79810	AbO79810	Pseudomon		184	17	50.0	1059	8	ADP31042	Adp31042	Human	sec
112	17	50.0	411	8	ADP30729	Adp30729	Human	sec	185	17	50.0	1065	8	ADP31347	Adp31347	Human	sec
113	17	50.0	412	8	ADT57242	Adt57242	Plant pol		186	17	50.0	1076	6	ABU81145	Abu81145	Human	PRO
114	17	50.0	420	8	ADP31273	Adp31273	Human	sec	187	17	50.0	1082	6	ABU66845	Abu66845	Human	PRO
115	17	50.0	420	8	ADP31274	Adp31274	Human	sec	188	17	50.0	1086	8	ADP30934	Adp30934	Human	sec
116	17	50.0	423	8	ADP31479	Adp31479	Human	sec	189	17	50.0	1092	8	ADP31358	Adp31358	Human	sec
117	17	50.0	483	8	ADP30852	Adp30852	Human	sec	190	17	50.0	1101	8	ADP31462	Adp31462	Human	sec
118	17	50.0	483	8	ADP30853	Adp30853	Human	sec	191	17	50.0	1110	8	ADP31430	Adp31430	Human	sec
119	17	50.0	510	8	ADP31049	Adp31049	Human	sec	192	17	50.0	1119	8	ADP31452	Adp31452	Human	sec
120	17	50.0	525	8	ADP31227	Adp31227	Human	sec	193	17	50.0	1128	8	ADP30483	Adp30483	Human	sec
121	17	50.0	531	8	ADP30529	Adp30529	Human	sec	194	17	50.0	1134	8	ADP30647	Adp30647	Human	sec
122	17	50.0	531	8	ADP31539	Adp31539	Human	sec	195	17	50.0	1134	8	ADP31537	Adp31537	Human	sec
123	17	50.0	531	8	ADP31540	Adp31540	Human	sec	196	17	50.0	1141	8	ADP30677	Adp30677	Human	sec
124	17	50.0	543	8	ADP30864	Adp30864	Human	sec	197	17	50.0	1170	8	ADP30919	Adp30919	Human	sec
125	17	50.0	552	8	ADP31019	Adp31019	Human	sec	198	17	50.0	1170	8	ADP30922	Adp30922	Human	sec
126	17	50.0	555	8	ADP31168	Adp31168	Human	sec	199	17	50.0	1173	8	ADP31155	Adp31155	Human	sec
127	17	50.0	559	8	ADP30716	Adp30716	Human	sec	200	17	50.0	1179	8	ADP30486	Adp30486	Human	sec
128	17	50.0	564	8	ADP30743	Adp30743	Human	sec	201	17	50.0	1189	8	ADP31043	Adp31043	Human	sec
129	17	50.0	567	8	ADP31565	Adp31565	Human	sec	202	17	50.0	1197	8	ADP31342	Adp31342	Human	sec
130	17	50.0	574	8	ADP31133	Adp31133	Human	sec	203	17	50.0	1205	8	ADP30935	Adp30935	Human	sec
131	17	50.0	592	8	ADP30917	Adp30917	Human	sec	204	17	50.0	1252	8	ADP30678	Adp30678	Human	sec
132	17	50.0	624	8	ADP31324	Adp31324	Human	sec	205	17	50.0	1358	8	ADP30995	Adp30995	Human	sec
133	17	50.0	624	8	ADP31325	Adp31325	Human	sec	206	17	50.0	1362	8	ADP31181	Adp31181	Human	sec
134	17	50.0	629	8	ADP31325	Adp31325	Human	sec	207	17	50.0	1383	8	ADP31091	Adp31091	Human	sec
135	17	50.0	645	8	ADP30858	Adp30858	Human	sec	208	17	50.0	1417	8	ADP31160	Adp31160	Human	sec
136	17	50.0	666	8	ADP31256	Adp31256	Human	sec	209	17	50.0	1420	8	ADP30944	Adp30944	Human	sec
137	17	50.0	681	8	ADP30598	Adp30598	Human	sec	210	17	50.0	1435	6	ABU09754	Abu09754	Human	hum
138	17	50.0	681	8	ADP30868	Adp30868	Human	sec	211	17	50.0	1435	6	ABU90136	Abu90136	Human	hum
139	17	50.0	702	8	ADP31518	Adp31518	Human	sec	212	17	50.0	1435	6	ABU96438	Abu96438	Human	hum
140	17	50.0	702	8	ADP31277	Adp31277	Human	sec	213	17	50.0	1435	6	ABU99047	Abu99047	Human	hum
141	17	50.0	702	8	ADP31278	Adp31278	Human	sec	214	17	50.0	1435	6	ABU98262	Abu98262	Human	hum
142	17	50.0	708	8	ADP31191	Adp31191	Human	sec	215	17	50.0	1435	6	ABU91968	Abu91968	Human	hum
143	17	50.0	711	8	ADP31535	Adp31535	Human	sec	216	17	50.0	1435	6	ABU08572	Abu08572	Human	hum
144	17	50.0	711	8	ADP31440	Adp31440	Human	sec	217	17	50.0	1435	6	ABU00411	Abu00411	Human	hum
145	17	50.0	714	8	ADP31561	Adp31561	Human	sec	218	17	50.0	1435	6	ABU88962	Abu88962	Human	hum
146	17	50.0	759	8	ADP31141	Adp31141	Human	sec	219	17	50.0	1435	6	ABU06458	Abu06458	Human	hum
147	17	50.0	768	8	ADP30912	Adp30912	Human	sec	220	17	50.0	1435	6	ABU95518	Abu95518	Human	hum
148	17	50.0	771	8	ADP30907	Adp30907	Human	sec	221	17	50.0	1435	6	ABU95208	Abu95208	Human	hum
149	17	50.0	774	8	ADP30885	Adp30885	Human	sec	222	17	50.0	1435	6	ABU90756	Abu90756	Human	hum
150	17	50.0	795	8	ADP31332	Adp31332	Human	sec	223	17	50.0	1435	6	ABU93918	Abu93918	Human	hum
151	17	50.0	799	8	ADP31261	Adp31261	Human	sec	224	17	50.0	1435	6	ABU86192	Abu86192	Human	hum
152	17	50.0	813	8	ADP30649	Adp30649	Human	sec	225	17	50.0	1435	6	ABU82047	Abu82047	Human	hum
153	17	50.0	834	8	ADP30738	Adp30738	Human	sec	226	17	50.0	1435	6	ABU97908	Abu97908	Human	hum
154	17	50.0	843	8	ADP30663	Adp30663	Human	sec	227	17	50.0	1435	6	ABU94428	Abu94428	Human	hum
155	17	50.0	848	8	ADP31037	Adp31037	Human	sec	228	17	50.0	1435	6	ABU00101	Abu00101	Human	hum
156	17	50.0	849	8	ADP31182	Adp31182	Human	sec	229	17	50.0	1435	6	ABU87112	Abu87112	Human	hum
157	17	50.0	849	8	ADP31275	Adp31275	Human	sec	230	17	50.0	1435	6	ABU91353	Abu91353	Human	hum
158	17	50.0	849	8	ADP31276	Adp31276	Human	sec	231	17	50.0	1435	6	ABU90446	Abu90446	Human	hum
159	17	50.0	861	8	ADP31340	Adp31340	Human	sec	232	17	50.0	1435	6	ABU97037	Abu97037	Human	hum
160	17	50.0	882	8	ADP30467	Adp30467	Human	sec	233	17	50.0	1435	6	ABU05233	Abu05233	Human	hum
161	17	50.0	895	8	ADJ92573	Adj92573	Human	PRO	234	17	50.0	1464	8	ADP31437	Adp31437	Human	sec
162	17	50.0	900	6	ADP31337	Adp31337	Human	sec	235	17	50.0	1470	8	ADP31611	Adp31611	Human	sec
163	17	50.0	908	6	ABU08402	Abu08402	Alpha-hel		236	17	50.0	1532	8	ADP31536	Adp31536	Human	sec
164	17	50.0	921	8	ADP31522	Adp31522	Human	sec	237	17	50.0	1550	8	ADP30567	Adp30567	Human	sec
165	17	50.0	930	8	ADP31444	Adp31444	Human	sec	238	17	50.0	1560	8	ADP31631	Adp31631	Human	sec
166	17	50.0	933	8	ADP31140	Adp31140	Human	sec	239	17	50.0	1614	8	ADP31529	Adp31529	Human	sec
167	17	50.0	939	8	ADP31542	Adp31542	Human	sec	240	17	50.0	1647	8	ADP30670	Adp30670	Human	sec
168	17	50.0	939	8	ADP31086	Adp31086	Human	sec	241	17	50.0	1647	8	ADP31419	Adp31419	Human	sec
169	17	50.0	939	8	ADP31541	Adp31541	Human	sec	242	17	50.0	1662	8	ADP31513	Adp31513	Human	sec
170	17	50.0	947	8	ADP30937	Adp30937	Human	sec	243	17	50.0	1700	3	ABU18144	Abu18144	Plaemodiu	

244	17	50.0	1743	6	ABU88255	Novel	hum	317	17	50.0	3390	8	ADP31148	Adp31148	Human	sec	
245	17	50.0	1743	6	ABU90134	Novel	hum	318	17	50.0	3447	8	ADP31112	Adp31112	Human	sec	
246	17	50.0	1743	6	ABU96436	Novel	hum	319	17	50.0	4315	5	ABP43908	Abp43908	MUC5B	par	
247	17	50.0	1743	6	ABU99045	Novel	hum	320	17	50.0	5304	8	ADP30706	Adp30706	Human	sec	
248	17	50.0	1743	6	ABU98260	Novel	hum	321	17	50.0	5314	8	ADP31186	Adp31186	Human	sec	
249	17	50.0	1743	6	ABU91966	Novel	hum	322	17	50.0	5514	8	ADP31591	Adp31591	Human	sec	
250	17	50.0	1743	6	ABU85270	Novel	hum	323	17	50.0	5703	8	ADL23265	Adl23265	Human	sec	
251	17	50.0	1743	6	ABO00409	Novel	hum	324	17	50.0	5858	9	ABE56507	Abe56507	Radiochem		
252	17	50.0	1743	6	ABU88960	Novel	hum	325	17	50.0	6465	8	ADP30705	Adp30705	Human	sec	
253	17	50.0	1743	6	ABO06456	Novel	hum	326	17	50.0	6729	8	ADP31600	Adp31600	Human	sec	
254	17	50.0	1743	6	ABU95516	Novel	hum	327	17	50.0	7285	6	ABJ38280	Abj38280	PAMG21-PA		
255	17	50.0	1743	6	ABU95206	Novel	hum	328	17	50.0	7339	6	AOI16358	Aoi16358	Human	tra	
256	17	50.0	1743	6	ABU90754	Novel	hum	329	17	50.0	10944	8	ADP31311	Adp31311	Human	sec	
257	17	50.0	1743	6	ABU93916	Novel	hum	330	17	50.0	10944	51	6	ABM53641	Abm53641	Proprionib	
258	17	50.0	1743	6	ABU86190	Novel	hum	331	17	50.0	11328	4	AAU59153	Aau59153	Proprionib		
259	17	50.0	1743	6	ABU82045	Novel	hum	332	17	50.0	52	6	ABM55672	Abm55672	Proprionib		
260	17	50.0	1743	6	ABU07906	Novel	hum	333	17	50.0	58	9	ADL11761	Adl11761	E. coli	D	
261	17	50.0	1743	6	ABU94226	Novel	hum	334	17	50.0	49	4	AAU57122	Aau57122	Proprionib		
262	17	50.0	1743	6	ABO00099	Novel	hum	335	17	50.0	51	6	ABM53641	Abm53641	Proprionib		
263	17	50.0	1743	6	ABU87110	Novel	hum	336	17	50.0	94	4	AAU59153	Aau59153	Proprionib		
264	17	50.0	1743	6	ABU91351	Novel	hum	337	17	50.0	52	6	ABM55672	Abm55672	Proprionib		
265	17	50.0	1743	6	ABU90444	Novel	hum	338	17	50.0	58	9	ADL11761	Adl11761	E. coli	D	
266	17	50.0	1743	6	ABU97035	Novel	hum	339	17	50.0	85	4	AAU61946	Aau61946	Proprionib		
267	17	50.0	1743	6	ABO05231	Novel	hum	340	17	50.0	85	6	ABM58465	Abm58465	Proprionib		
268	17	50.0	1776	9	ADY62715	Human	alp	341	17	50.0	85	6	ABM58465	Abm58465	Proprionib		
269	17	50.0	1782	8	ADP31270	Human	sec	342	17	50.0	100	8	ADP65826	Adp65826	Plant	ful	
270	17	50.0	1803	8	ADP30562	Human	sec	343	17	50.0	109	8	ADP30690	Adp30690	Human	sec	
271	17	50.0	1894	6	ABU88253	Novel	hum	344	17	50.0	132	7	ADP94844	Adp94844	Generic	b	
272	17	50.0	1894	6	ABU90132	Novel	hum	345	17	50.0	133	7	ABO70871	AbO70871	Pseudomon		
273	17	50.0	1894	6	ABU96434	Novel	hum	346	17	50.0	134	8	ABO94811	AbO94811	Generic	b	
274	17	50.0	1894	6	ABU99043	Novel	hum	347	17	50.0	138	7	ABO67837	AbO67837	Pseudomon		
275	17	50.0	1894	6	ABU98258	Novel	hum	348	17	50.0	142	7	ABO75180	AbO75180	Pseudomon		
276	17	50.0	1894	6	ABU91964	Novel	hum	349	17	50.0	144	7	ABO77101	AbO77101	Pseudomon		
277	17	50.0	1894	6	ABU85268	Novel	hum	350	17	50.0	144	8	ADP72498	Adp72498	Plant	ful	
278	17	50.0	1894	6	ABO00407	Novel	hum	351	17	50.0	151	7	ABO81297	AbO81297	Pseudomon		
279	17	50.0	1894	6	ABU88958	Novel	hum	352	17	50.0	154	7	ABO80026	AbO80026	Pseudomon		
280	17	50.0	1894	6	ABO06454	Novel	hum	353	17	50.0	155	7	ABO71527	AbO71527	Pseudomon		
281	17	50.0	1894	6	ABU95514	Novel	hum	354	17	50.0	159	8	ADP30763	Adp30763	Human	sec	
282	17	50.0	1894	6	ABU95204	Novel	hum	355	17	50.0	159	8	ADP30762	Adp30762	Human	sec	
283	17	50.0	1894	6	ABU90752	Novel	hum	356	17	50.0	160	4	AAU95024	Aau95024	Human	rep	
284	17	50.0	1894	6	ABU93914	Novel	hum	357	17	50.0	160	4	ABM95732	Abm95732	Human	tes	
285	17	50.0	1894	6	ABU86188	Novel	hum	358	17	50.0	165	8	ADP31286	Adp31286	Human	sec	
286	17	50.0	1894	6	ABU82043	Novel	hum	359	17	50.0	165	8	ADP31174	Adp31174	Human	sec	
287	17	50.0	1894	6	ABU07904	Novel	hum	360	17	50.0	171	8	ADP30794	Adp30794	Human	sec	
288	17	50.0	1894	6	ABU94224	Novel	hum	361	17	50.0	171	8	ADP31497	Adp31497	Human	sec	
289	17	50.0	1894	6	ABO00097	Novel	hum	362	17	50.0	171	8	ADP30793	Adp30793	Human	sec	
290	17	50.0	1894	6	ABU87108	Novel	hum	363	17	50.0	180	8	ADP30820	Adp30820	Human	sec	
291	17	50.0	1894	6	ABU91349	Novel	hum	364	17	50.0	180	8	ADP30825	Adp30825	Human	sec	
292	17	50.0	1894	6	ABU90442	Novel	hum	365	17	50.0	180	8	ADP30828	Adp30828	Human	sec	
293	17	50.0	1894	6	ABU97033	Novel	hum	366	17	50.0	180	8	ADP30826	Adp30826	Human	sec	
294	17	50.0	1894	6	ABO05229	Novel	hum	367	17	50.0	180	8	ADP30827	Adp30827	Human	sec	
295	17	50.0	1894	6	ADP30669	Human	sec	368	17	50.0	180	8	ADP31660	Adp31660	Human	sec	
296	17	50.0	2058	8	ADP31630	Human	sec	369	17	50.0	188	8	ADP31660	Adp31660	Human	sec	
297	17	50.0	2091	8	ADP31088	Human	sec	370	17	50.0	195	8	ADP30590	Adp30590	Human	sec	
298	17	50.0	2109	8	ADP07505	Human	col	371	17	50.0	198	2	AAU59841	Aau59841	Human	sec	
299	17	50.0	2124	8	ADP30560	Human	col	372	17	50.0	198	2	AAU59841	Aau59841	Human	sec	
300	17	50.0	2187	8	ADP30644	Human	sec	373	17	50.0	198	8	ADP30766	Adp30766	Human	sec	
301	17	50.0	2254	8	ADP30882	Human	sec	374	17	50.0	198	8	ADP30767	Adp30767	Human	sec	
302	17	50.0	2254	8	ADP30882	Human	sec	375	17	50.0	199	7	ABO73562	AbO73562	Human	sec	
303	17	50.0	2260	8	ADP30687	Human	col	376	17	50.0	201	1	ADP30711	Adp30711	Human	sec	
304	17	50.0	2272	8	ADP31136	Human	sec	377	17	50.0	215	9	ABE803544	Abe803544	MycoBacte		
305	17	50.0	2304	8	ADP30669	Human	sec	378	17	50.0	215	9	ABE79447	Abe79447	Human	sec	
306	17	50.0	2304	8	ADP31252	Human	sec	379	17	50.0	217	7	ABO72885	AbO72885	Human	sec	
307	17	50.0	2307	8	ADP31334	Human	sec	380	17	50.0	219	7	ABO76425	AbO76425	Pseudomon		
308	17	50.0	2349	8	ADP30959	Human	sec	381	17	50.0	225	8	ADP30589	Adp30589	Human	sec	
309	17	50.0	2382	8	ADP31341	Human	sec	382	17	50.0	228	8	ADP31281	Adp31281	Human	sec	
310	17	50.0	2401	8	ADP07507	Human	col	383	17	50.0	229	3	ADP07930	Adp07930	Rice	prot	
311	17	50.0	2415	8	ADP31023	Human	col	384	17	50.0	234	3	AAU93028	Aau93028	Arabidops		
312	17	50.0	2418	8	ADP31105	Human	sec	385	17	50.0	234	5	AAU93028	Aau93028	Arabidops		
313	17	50.0	2616	8	ADP31253	Human	sec	386	17	50.0	234	5	ADP30781	Adp30781	Plant	yie	
314	17	50.0	2616	9	ABE84965	N. mening		387	17	50.0	234	7	ADP30781	Adp30781	Plant	tra	
315	17	50.0	2976	8	ADP30724	Human	sec	388	17	50.0	234	8	ADP30724	Adp30724	Human	sec	
316	17	50.0	3339	8	ADP31219	Human	sec	389	17	50.0	234	8	ADP31219	Adp31219	Human	sec	

390	16	47.1	234	8	AD002129	Thalectes	463	16	47.1	471	8	ADP30854	Adp30854	Human	sec
391	16	47.1	234	8	ADP30484	Human	464	16	47.1	472	8	ADP31222	Adp31222	Human	sec
392	16	47.1	234	8	ADV09559	A. thalia	465	16	47.1	474	8	ADP31083	Adp31083	Human	sec
393	16	47.1	234	8	ADV09550	A. thalia	466	16	47.1	477	8	ADP31235	Adp31235	Human	sec
394	16	47.1	243	8	ADP31343	Human	467	16	47.1	480	8	ADP31484	Adp31484	Human	sec
395	16	47.1	246	8	ADP30619	Human	468	16	47.1	489	8	ADP31380	Adp31380	Human	sec
396	16	47.1	249	7	AB069997	Pseudomon	469	16	47.1	493	9	ADW69641	Adw69641	HIV-1	SP1
397	16	47.1	254	8	ADP30739	Human	470	16	47.1	501	8	ADP31689	Adp31689	Human	sec
398	16	47.1	272	8	ADP30633	Human	471	16	47.1	510	8	ADP31549	Adp31549	Human	sec
399	16	47.1	278	8	ADQ66574	Novel hum	472	16	47.1	518	8	ADP31190	Adp31190	Human	sec
400	16	47.1	279	8	ADP31523	Human	473	16	47.1	529	8	ADP31525	Adp31525	Human	sec
401	16	47.1	280	8	ADP31489	Human	474	16	47.1	531	8	ADP31696	Adp31696	Human	sec
402	16	47.1	280	7	AB072037	Pseudomon	475	16	47.1	533	8	ADP31279	Adp31279	Human	sec
403	16	47.1	285	7	AB072271	Pseudomon	476	16	47.1	543	8	ADP31329	Adp31329	Human	sec
404	16	47.1	286	8	ADP31685	Human	477	16	47.1	546	8	ADP31449	Adp31449	Human	sec
405	16	47.1	291	8	ADP31193	Human	478	16	47.1	549	8	ADP30855	Adp30855	Human	sec
406	16	47.1	294	8	ADP31045	Human	479	16	47.1	549	8	ADP31639	Adp31639	Human	sec
407	16	47.1	306	7	AB074423	Pseudomon	480	16	47.1	549	8	ADP31009	Adp31009	Human	sec
408	16	47.1	311	2	AAR59843	ApobFlx2	481	16	47.1	552	8	ADP31524	Adp31524	Human	sec
409	16	47.1	311	2	AAR92115	Human	482	16	47.1	557	8	ADP31103	Adp31103	Human	sec
410	16	47.1	315	8	ADP31685	Human	483	16	47.1	558	8	ADP31254	Adp31254	Human	sec
411	16	47.1	317	8	ADP31615	Human	484	16	47.1	558	8	ADP31255	Adp31255	Human	sec
412	16	47.1	318	8	ADP30824	Human	485	16	47.1	563	8	ADP30881	Adp30881	Human	sec
413	16	47.1	318	8	ADP31135	Human	486	16	47.1	585	8	ADP31458	Adp31458	Human	sec
414	16	47.1	320	7	AB075888	Pseudomon	487	16	47.1	600	8	ADP30865	Adp30865	Human	sec
415	16	47.1	331	7	AB070026	Pseudomon	488	16	47.1	604	8	ADP30940	Adp30940	Human	sec
416	16	47.1	333	8	ADP30848	Human	489	16	47.1	604	8	ADP30941	Adp30941	Human	sec
417	16	47.1	336	8	ADP31320	Human	490	16	47.1	605	8	ADL12887	Adl12887	Human	sec
418	16	47.1	345	8	ADP31016	Human	491	16	47.1	610	8	ADP31264	Adp31264	Human	sec
419	16	47.1	346	8	ADP31355	Human	492	16	47.1	612	8	ADP31169	Adp31169	Human	sec
420	16	47.1	346	8	ADK68092	Plant	493	16	47.1	618	8	ADP31467	Adp31467	Human	sec
421	16	47.1	348	8	ADP30985	Human	494	16	47.1	618	8	ADP31466	Adp31466	Human	sec
422	16	47.1	348	8	ADP31676	Human	495	16	47.1	626	8	ADP30528	Adp30528	Human	sec
423	16	47.1	354	8	ADP30783	Human	496	16	47.1	627	8	ADP31388	Adp31388	Human	sec
424	16	47.1	354	8	ADP30707	Human	497	16	47.1	629	7	AB068031	Ab068031	Pseudomon	
425	16	47.1	354	8	ADP30779	Human	498	16	47.1	637	8	ADP30539	Adp30539	Human	sec
426	16	47.1	354	8	ADP30784	Human	499	16	47.1	639	8	ADP30861	Adp30861	Human	sec
427	16	47.1	354	8	ADP30778	Human	500	16	47.1	642	8	ADP31265	Adp31265	Human	sec
428	16	47.1	357	8	ADP31560	Human	501	16	47.1	642	8	ADP30748	Adp30748	Human	sec
429	16	47.1	357	8	ADP30505	Human	502	16	47.1	642	8	ADP31161	Adp31161	Human	sec
430	16	47.1	360	8	ADP31439	Human	503	16	47.1	645	8	ADP31124	Adp31124	Human	sec
431	16	47.1	363	8	ADP30659	Human	504	16	47.1	645	8	ADP31125	Adp31125	Human	sec
432	16	47.1	369	8	ADP30807	Human	505	16	47.1	645	8	ADP31183	Adp31183	Human	sec
433	16	47.1	372	8	ADP31492	Human	506	16	47.1	651	8	ADP31022	Adp31022	Human	sec
434	16	47.1	372	8	ADP31610	Human	507	16	47.1	658	8	ADP31226	Adp31226	Human	sec
435	16	47.1	378	8	ADP31375	Human	508	16	47.1	660	8	ADP31606	Adp31606	Human	sec
436	16	47.1	381	8	ADP30580	Human	509	16	47.1	665	8	ADP30571	Adp30571	Human	sec
437	16	47.1	382	8	ADP31221	Human	510	16	47.1	667	8	ADP30867	Adp30867	Human	sec
438	16	47.1	384	8	ADP31338	Human	511	16	47.1	667	8	ADP04463	Adp04463	Sea	squid
439	16	47.1	387	8	ADP31368	Human	512	16	47.1	669	8	ADP31598	Adp31598	Human	sec
440	16	47.1	387	8	ADP31377	Human	513	16	47.1	672	8	ADP31667	Adp31667	Human	sec
441	16	47.1	392	8	ADT59578	Plant	514	16	47.1	672	8	ADP31666	Adp31666	Human	sec
442	16	47.1	396	8	ADP31378	Human	515	16	47.1	678	8	ADP31585	Adp31585	Human	sec
443	16	47.1	399	8	ADP31464	Human	516	16	47.1	678	8	ADP31283	Adp31283	Human	sec
444	16	47.1	411	7	ABM67022	Rice	517	16	47.1	681	8	ADP31053	Adp31053	Human	sec
445	16	47.1	411	8	ADP31104	Human	518	16	47.1	690	8	ADP30750	Adp30750	Human	sec
446	16	47.1	414	8	ADP31477	Human	519	16	47.1	699	8	ADP30747	Adp30747	Human	sec
447	16	47.1	420	8	ADP31349	Human	520	16	47.1	705	8	ADP31330	Adp31330	Human	sec
448	16	47.1	421	8	ADP31159	Human	521	16	47.1	705	8	ADP31623	Adp31623	Human	sec
449	16	47.1	426	8	ADP31331	Human	522	16	47.1	710	8	ADP30746	Adp30746	Human	sec
450	16	47.1	426	8	ADP31495	Human	523	16	47.1	711	8	ADP31652	Adp31652	Human	sec
451	16	47.1	429	8	ADP30570	Human	524	16	47.1	728	8	ADP30508	Adp30508	Human	sec
452	16	47.1	429	8	ADP31207	Human	525	16	47.1	729	8	ADP30541	Adp30541	Human	sec
453	16	47.1	435	8	ADP31102	Human	526	16	47.1	746	8	ADP30982	Adp30982	Human	sec
454	16	47.1	438	8	ADP30818	Human	527	16	47.1	750	8	ADP30576	Adp30576	Human	sec
455	16	47.1	441	8	ADP31206	Human	528	16	47.1	750	8	ADP30857	Adp30857	Human	sec
456	16	47.1	444	8	ADP31640	Human	529	16	47.1	750	8	ADP31131	Adp31131	Human	sec
457	16	47.1	449	8	ADQ21283	Human	530	16	47.1	750	8	ADP30856	Adp30856	Human	sec
458	16	47.1	450	8	ADP31085	Human	531	16	47.1	764	8	ADP31614	Adp31614	Human	sec
459	16	47.1	453	8	ADP31465	Human	532	16	47.1	768	8	ADP31126	Adp31126	Human	sec
460	16	47.1	459	8	ADP30932	Human	533	16	47.1	772	8	ADP30936	Adp30936	Human	sec
461	16	47.1	459	8	ADP30930	Human	534	16	47.1	774	8	ADP30506	Adp30506	Human	sec
462	16	47.1	467	8	ADX31801	Plant	535	16	47.1	774	8	ADP31373	Adp31373	Human	sec

536	16	47.1	774	8	ADP30780	Human sec	609	16	47.1	1221	8	ADP30884	Adp30884	Human sec
537	16	47.1	774	8	ADP30781	Human sec	610	16	47.1	1224	8	ADP31426	Adp31426	Human sec
538	16	47.1	774	8	ADP31235	Human sec	611	16	47.1	1227	8	ADP31602	Adp31602	Human sec
539	16	47.1	776	7	ADP71157	Human int	612	16	47.1	1227	8	ADP31210	Adp31210	Human sec
540	16	47.1	776	8	ABM82451	Human int	613	16	47.1	1233	8	ADP30523	Adp30523	Human sec
541	16	47.1	776	9	ADY16233	Humour-as	614	16	47.1	1239	8	ADP31297	Adp31297	Human sec
542	16	47.1	783	8	ADP31284	Human sec	615	16	47.1	1260	8	ADP30883	Adp30883	Human sec
543	16	47.1	783	8	ADP31436	Human sec	616	16	47.1	1260	8	ADP31189	Adp31189	Human sec
544	16	47.1	792	9	AEA20901	Novel hum	617	16	47.1	1269	8	ADU23456	Adu23456	Sulfolobu
545	16	47.1	804	8	ADP31635	Human sec	618	16	47.1	1282	8	ADP31328	Adp31328	Human sec
546	16	47.1	807	8	ADP31036	Human sec	619	16	47.1	1302	8	ADP31595	Adp31595	Human sec
547	16	47.1	813	8	ADP31282	Human sec	620	16	47.1	1305	8	ADP31139	Adp31139	Human sec
548	16	47.1	813	8	ADP30561	Human sec	621	16	47.1	1314	8	ADP31197	Adp31197	Human sec
549	16	47.1	816	8	ADP31245	Human sec	622	16	47.1	1344	8	ADP31211	Adp31211	Human sec
550	16	47.1	816	8	ADP31581	Human sec	623	16	47.1	1350	8	ADP31138	Adp31138	Human sec
551	16	47.1	828	8	ADP30878	Human sec	624	16	47.1	1359	8	ADP31129	Adp31129	Human sec
552	16	47.1	831	8	ADP31179	Human sec	625	16	47.1	1365	8	ADP31561	Adp31561	Human sec
553	16	47.1	852	8	ADP30664	Human sec	626	16	47.1	1371	8	ADP31646	Adp31646	Human sec
554	16	47.1	861	8	ADP31020	Human sec	627	16	47.1	1371	8	ADP30876	Adp30876	Human sec
555	16	47.1	861	8	ADP31021	Human sec	628	16	47.1	1380	8	ADP31566	Adp31566	Human sec
556	16	47.1	885	8	ADP31198	Human sec	629	16	47.1	1386	8	ADP31371	Adp31371	Human sec
557	16	47.1	887	8	ADP30554	Human sec	630	16	47.1	1387	8	ADP30946	Adp30946	Human sec
558	16	47.1	887	8	ADP30558	Human sec	631	16	47.1	1431	8	ADP31609	Adp31609	Human sec
559	16	47.1	888	8	ADP30971	Human sec	632	16	47.1	1440	8	ADP34533	Adp34533	Glucocort
560	16	47.1	890	8	ADP31059	Human sec	633	16	47.1	1456	8	ADP30923	Adp30923	Human sec
561	16	47.1	891	8	ADP31668	Human sec	634	16	47.1	1464	7	ABO80681	AbO80681	Pseudomon
562	16	47.1	897	8	ADP30914	Human sec	635	16	47.1	1476	8	ADP31550	Adp31550	Human sec
563	16	47.1	912	8	ADP31507	Human sec	636	16	47.1	1494	8	ADP31560	Adp31560	Human sec
564	16	47.1	918	8	ADP31459	Human sec	637	16	47.1	1506	8	ADP30596	Adp30596	Human sec
565	16	47.1	933	8	ADP31510	Human sec	638	16	47.1	1518	8	ADP31532	Adp31532	Human sec
566	16	47.1	936	8	ADP31486	Human sec	639	16	47.1	1566	8	ADP31407	Adp31407	Human sec
567	16	47.1	936	8	ADP31597	Human sec	640	16	47.1	1578	8	ADP30499	Adp30499	Human sec
568	16	47.1	939	8	ADP30726	Human sec	641	16	47.1	1584	8	ADP31405	Adp31405	Human sec
569	16	47.1	945	8	ADP31238	Human sec	642	16	47.1	1596	8	ADP31491	Adp31491	Human sec
570	16	47.1	945	8	ADP31237	Human sec	643	16	47.1	1602	9	ADZ66459	AdZ66459	Human ABN
571	16	47.1	948	8	ADP30586	Human sec	644	16	47.1	1623	8	ADP30552	Adp30552	Human sec
572	16	47.1	966	8	ADP30745	Human sec	645	16	47.1	1623	8	ADP31653	Adp31653	Human sec
573	16	47.1	1001	8	ADP31356	Human sec	646	16	47.1	1629	8	ADP30945	Adp30945	Human sec
574	16	47.1	1002	8	ADP31369	Human sec	647	16	47.1	1647	8	ADP31052	Adp31052	Human sec
575	16	47.1	1002	8	ADP30866	Human pan	648	16	47.1	1670	8	ADP30927	Adp30927	Human sec
576	16	47.1	1008	8	ADP30721	Human sec	649	16	47.1	1719	8	ADP31137	Adp31137	Human sec
577	16	47.1	1010	8	ADP31296	Human sec	650	16	47.1	1746	8	ADP30992	Adp30992	Human sec
578	16	47.1	1017	6	AAO26719	SR protel	651	16	47.1	1749	8	ADP31408	Adp31408	Human sec
579	16	47.1	1017	8	ADP31268	Human sec	652	16	47.1	1771	8	ADP30674	Adp30674	Human sec
580	16	47.1	1023	8	ADP30994	Human sec	653	16	47.1	1782	8	ADP31391	Adp31391	Human sec
581	16	47.1	1030	8	ADP30913	Human sec	654	16	47.1	1789	8	ADP30962	Adp30962	Human sec
582	16	47.1	1050	8	ADP31370	Human sec	655	16	47.1	1794	8	ADP31176	Adp31176	Human sec
583	16	47.1	1056	8	ADP31082	Human sec	656	16	47.1	1815	8	ADP31601	Adp31601	Human sec
584	16	47.1	1057	8	ADP31592	Human sec	657	16	47.1	1827	8	ADP31170	Adp31170	Human sec
585	16	47.1	1065	8	ADP31369	Human sec	658	16	47.1	1845	8	ADP31372	Adp31372	Human sec
586	16	47.1	1065	8	ADP30661	Human sec	659	16	47.1	1875	8	ADP31664	Adp31664	Human sec
587	16	47.1	1065	8	ADP30973	Human sec	660	16	47.1	1878	8	ADP30740	Adp30740	Human sec
588	16	47.1	1065	8	ADP31482	Human sec	661	16	47.1	1913	6	ADA15719	AdA15719	C. elegan
589	16	47.1	1065	8	ADP31287	Human sec	662	16	47.1	1937	8	ADP30889	Adp30889	Human sec
590	16	47.1	1066	8	ADP30563	Human sec	663	16	47.1	1933	8	ADP30902	Adp30902	Human sec
591	16	47.1	1086	8	ADP31447	Human sec	664	16	47.1	1956	8	ADP31662	Adp31662	Human sec
592	16	47.1	1090	8	ADQ10187	Human pol	665	16	47.1	1989	8	ADP31185	Adp31185	Human sec
593	16	47.1	1104	8	ADP31156	Human sec	666	16	47.1	1996	8	ADP31599	Adp31599	Human sec
594	16	47.1	1116	8	ADP31128	Human sec	667	16	47.1	2001	8	ADP31644	Adp31644	Human sec
595	16	47.1	1132	8	ADP30952	Human sec	668	16	47.1	2020	8	ADP31056	Adp31056	Human sec
596	16	47.1	1140	8	ADP31128	Human sec	669	16	47.1	2148	8	ADP30974	Adp30974	Human sec
597	16	47.1	1140	8	ADP31130	Human sec	670	16	47.1	2358	8	ADP31690	Adp31690	Human sec
598	16	47.1	1147	8	ADP30965	Human sec	671	16	47.1	2391	8	ADP31366	Adp31366	Human sec
599	16	47.1	1183	8	ADP30556	Human sec	672	16	47.1	2454	8	ADP30469	Adp30469	Human sec
600	16	47.1	1184	7	AAO6553	Banana CK	673	16	47.1	2469	8	ADP31031	Adp31031	Human sec
601	16	47.1	1186	2	AAO5840	Banana ri	674	16	47.1	2514	8	ADP30727	Adp30727	Human sec
602	16	47.1	1194	8	ADP31146	Human sec	675	16	47.1	2535	8	ADP31146	Adp31146	Human sec
603	16	47.1	1199	8	ADP31044	Human sec	676	16	47.1	2547	8	ADP31665	Adp31665	Human sec
604	16	47.1	1200	8	ADP31603	Human sec	677	16	47.1	2611	8	ADP31027	Adp31027	Human sec
605	16	47.1	1200	9	ADZ76047	Mouse pre	678	16	47.1	2622	8	ADP31663	Adp31663	Human sec
606	16	47.1	1209	8	ADP30782	Human sec	679	16	47.1	2664	8	ADP31546	Adp31546	Human sec
607	16	47.1	1215	8	ADP30898	Human sec	680	16	47.1	2828	8	ADP30938	Adp30938	Human sec
608	16	47.1	1215	8	ADP30900	Human sec	681	16	47.1	2980	9	AEA36049	AdA36049	Maize Sta

682	16	47.1	3036	8	ADP31595	Adp31595	Human sec	755	15	44.1	111	7	ABO83484	ABO83484	Pseudomon
683	16	47.1	3046	8	AD010419	Adg10419	Human pol	756	15	44.1	112	5	ABB89458	ABB89458	Human pol
684	16	47.1	3144	8	ADP31544	Adp31544	Human sec	757	15	44.1	113	8	ADP31517	ADP31517	Human sec
685	16	47.1	3201	8	ADP31545	Adp31545	Human sec	758	15	44.1	118	4	AAU53541	AAU53541	Proprionib
686	16	47.1	3465	8	ADP31234	Adp31234	Human sec	759	15	44.1	118	6	ABM50060	ABM50060	Pseudomon
687	16	47.1	3477	8	ADP30704	Adp30704	Human sec	760	15	44.1	123	7	ABO70463	ABO70463	Pseudomon
688	16	47.1	3579	8	ADP31098	Adp31098	Human sec	761	15	44.1	125	4	AAU47488	AAU47488	Proprionib
689	16	47.1	3579	8	ADP30991	Adp30991	Human sec	762	15	44.1	125	6	ABM44007	ABM44007	Proprionib
690	16	47.1	3507	5	ABG70822	Abg70822	Mouse myo	763	15	44.1	126	6	ABR57124	ABR57124	Camel ant
691	16	47.1	3507	6	ABG74190	Abg74190	Mouse myo	764	15	44.1	127	8	ADP31622	ADP31622	Human sec
692	16	47.1	4752	8	ADP30585	Adp30585	Human sec	765	15	44.1	130	7	ADP27051	ADP27051	Pseudomon
693	16	47.1	4752	8	ADP30651	Adp30651	Human sec	766	15	44.1	131	7	ABO66182	ABO66182	Human adi
694	16	47.1	5397	8	ADP31068	Adp31068	Human sec	767	15	44.1	134	7	ABO76972	ABO76972	Pseudomon
695	16	47.1	8973	8	ADP31119	Adp31119	Human sec	768	15	44.1	135	7	ABO83194	ABO83194	Pseudomon
696	15	44.1	12	6	ABG71386	Abg71386	Staphyloc	769	15	44.1	135	7	ABO79475	ABO79475	Pseudomon
697	15	44.1	12	6	ADP95109	Adp95109	PCR prime	770	15	44.1	135	8	ADP30703	ADP30703	Human sec
698	15	44.1	24	8	ADU59677	Adus9677	GBP-4 c10	771	15	44.1	136	7	ABO74222	ABO74222	Pseudomon
699	15	44.1	27	4	AAU13578	Aam13578	Human bra	772	15	44.1	138	3	AAU42027	AAU42027	Human ORF
700	15	44.1	27	4	ABB32506	Abb32506	Peptide #	773	15	44.1	139	7	ABO68581	ABO68581	Pseudomon
701	15	44.1	27	4	AAU25975	Aam25975	Peptide #	774	15	44.1	142	8	ADP31632	ADP31632	Human sec
702	15	44.1	27	4	ABR27361	Abb27361	Human pep	775	15	44.1	143	7	ABO77067	ABO77067	Pseudomon
703	15	44.1	27	4	ABR18013	Abb18013	Protein #	776	15	44.1	144	7	ABO75730	ABO75730	Pseudomon
704	15	44.1	27	4	AAU53341	Aam53341	Human bon	777	15	44.1	146	4	ABG02814	ABG02814	Novel hum
705	15	44.1	27	4	AAU53341	Aam53341	Human bra	778	15	44.1	150	8	AD145261	AD145261	Human CGD
706	15	44.1	27	4	ABG47359	Abg47359	Human liv	779	15	44.1	152	8	ADP31605	ADP31605	Human sec
707	15	44.1	27	4	AAU01330	Aam01330	Peptide #	780	15	44.1	153	7	ABO83454	ABO83454	Pseudomon
708	15	44.1	27	5	ABG35348	Abg35348	Human pep	781	15	44.1	154	7	ABO75839	ABO75839	Pseudomon
709	15	44.1	34	4	AAU17951	Aam17951	Peptide #	782	15	44.1	156	7	ABO75667	ABO75667	Pseudomon
710	15	44.1	34	4	ABR36982	Abb36982	Peptide #	783	15	44.1	158	3	AAU29118	AAU29118	Pseudomon
711	15	44.1	41	8	ADP21826	Adp21826	Low dens1	784	15	44.1	163	7	ADP78226	ADP78226	Human sec
712	15	44.1	61	4	AAU51383	Aau51383	Proprionib	785	15	44.1	163	7	ABO80383	ABO80383	Pseudomon
713	15	44.1	61	6	ABM47902	Abm47902	Proprionib	786	15	44.1	165	8	ADP30587	ADP30587	Human sec
714	15	44.1	69	4	AAU62232	Aau62232	Proprionib	787	15	44.1	165	8	ADP31041	ADP31041	Human sec
715	15	44.1	69	6	ABM58751	Abm58751	Proprionib	788	15	44.1	168	8	ADP30841	ADP30841	Human sec
716	15	44.1	70	4	ABR04032	Abb04032	Human mus	789	15	44.1	169	7	ABO74424	ABO74424	Pseudomon
717	15	44.1	70	6	ADU329352	Adu329352	Novel hum	790	15	44.1	171	8	ADP30599	ADP30599	Human sec
718	15	44.1	72	7	ABO66669	Abb66669	Klebsiell	791	15	44.1	171	8	ADY10989	ADY10989	Plant ful
719	15	44.1	72	7	ABR37627	Abb37627	Peptide #	792	15	44.1	173	8	ADP31530	ADP31530	Human sec
720	15	44.1	76	4	AAU70732	Aam70732	Human bon	793	15	44.1	177	7	ABO70036	ABO70036	Pseudomon
721	15	44.1	76	4	ABG52439	Abg52439	Human liv	794	15	44.1	177	8	ADP30755	ADP30755	Human sec
722	15	44.1	76	4	AAU06152	Aam06152	Peptide #	795	15	44.1	179	7	ABO78051	ABO78051	Pseudomon
723	15	44.1	83	6	ABM49684	Abm49684	Proprionib	796	15	44.1	187	8	AAU54675	AAU54675	Pseudomon
724	15	44.1	85	4	AAU46868	Aau46868	Proprionib	797	15	44.1	183	4	ABM51194	ABM51194	Proprionib
725	15	44.1	80	4	AAU59349	Aau59349	Proprionib	798	15	44.1	183	6	ABO77341	ABO77341	Pseudomon
726	15	44.1	80	6	ABM55868	Abm55868	Proprionib	799	15	44.1	184	7	ABE803464	ABE803464	Mycobacte
727	15	44.1	80	7	ABO71261	AbO71261	Pseudomon	800	15	44.1	187	9	AEA79367	AEA79367	Novel M.
728	15	44.1	83	4	AAU53165	Aau53165	Proprionib	801	15	44.1	187	9	ABM69571	ABM69571	Phototrab
729	15	44.1	83	6	ABM49684	Abm49684	Proprionib	802	15	44.1	188	6	ABO81563	ABO81563	Pseudomon
730	15	44.1	85	4	AAU46868	Aau46868	Proprionib	803	15	44.1	188	7	ADP31641	ADP31641	Human sec
731	15	44.1	85	6	ABM43387	Abm43387	Proprionib	804	15	44.1	189	8	ADP30641	ADP30641	Human sec
732	15	44.1	86	7	ABO78269	AbO78269	Pseudomon	805	15	44.1	189	8	ADP30641	ADP30641	Human sec
733	15	44.1	87	4	AAU41562	Aau41562	Proprionib	806	15	44.1	192	7	AD163118	AD163118	Human apo
734	15	44.1	87	4	ABG07613	Abg07613	Novel hum	807	15	44.1	192	8	ADP30573	ADP30573	Human sec
735	15	44.1	87	6	ABM38081	Abm38081	Proprionib	808	15	44.1	192	8	ADP30575	ADP30575	Human sec
736	15	44.1	88	7	ABO76533	AbO76533	Pseudomon	809	15	44.1	192	8	ADP30492	ADP30492	Human sec
737	15	44.1	89	4	ABG27446	Abg27446	Novel hum	810	15	44.1	194	7	ABO78166	ABO78166	Pseudomon
738	15	44.1	90	6	ADP34228	Adp34228	Acinetoba	811	15	44.1	194	8	ADP31626	ADP31626	Human sec
739	15	44.1	95	4	AAU50672	Aau50672	Proprionib	812	15	44.1	195	8	ADP30696	ADP30696	Human sec
740	15	44.1	95	4	AAU40381	Aau40381	Proprionib	813	15	44.1	198	8	ADP30493	ADP30493	Human sec
741	15	44.1	95	6	ABM47191	Abm47191	Proprionib	814	15	44.1	198	8	ADP30477	ADP30477	Human sec
742	15	44.1	95	6	ABM36900	Abm36900	Proprionib	815	15	44.1	198	8	ADP30477	ADP30477	Human sec
743	15	44.1	101	4	AAU83737	Aau83737	Human imm	816	15	44.1	198	8	ADP30481	ADP30481	Human sec
744	15	44.1	102	4	AAU61573	Abm61573	Proprionib	817	15	44.1	200	7	ABO73557	ABO73557	Pseudomon
745	15	44.1	102	6	ABM58092	Abm58092	Proprionib	818	15	44.1	200	7	ABO73557	ABO73557	Pseudomon
746	15	44.1	103	6	AAU48964	Aau48964	Proprionib	819	15	44.1	201	3	AAU29117	AAU29117	Arabidops
747	15	44.1	103	6	ABM45503	Abm45503	Proprionib	820	15	44.1	204	8	ADP30488	ADP30488	Human sec
748	15	44.1	103	7	ABO84304	AbO84304	Pseudomon	821	15	44.1	204	8	ADP31421	ADP31421	Human sec
749	15	44.1	106	2	AAU29187	Aau29187	Amino aci	822	15	44.1	204	8	ADP31422	ADP31422	Human sec
750	15	44.1	107	4	AAU63314	Aau63314	Proprionib	823	15	44.1	204	8	ADP30489	ADP30489	Human sec
751	15	44.1	107	5	ABP64460	Abp64460	Human ORF	824	15	44.1	204	8	ADP31424	ADP31424	Human sec
752	15	44.1	107	6	ABM39833	Abm39833	Proprionib	825	15	44.1	204	8	ADP31420	ADP31420	Human sec
753	15	44.1	108	3	AAU21190	Aau21190	Exo14 par	826	15	44.1	204	8	ADP30490	ADP30490	Human sec
754	15	44.1	110	9	ABM95190	Abm95190	M. xanthu	827	15	44.1	204	8	ADP30490	ADP30490	Human sec

828	15	44.1	207	7	ABO71507	Pseudomon	901	15	44.1	339	8	ADP30519	Adp30519	Human sec
829	15	44.1	208	7	ABO75819	Pseudomon	902	15	44.1	339	8	ADP30702	Adp30702	Human sec
830	15	44.1	208	7	ABO78684	Pseudomon	903	15	44.1	348	8	ADP31374	Adp31374	Human sec
831	15	44.1	210	8	ADP30986	Human sec	904	15	44.1	349	5	ABO80605	ABO80605	Hirudin P
832	15	44.1	210	8	ADP30960	Human sec	905	15	44.1	350	4	ABO86835	ABO86835	Drosophil
833	15	44.1	212	2	AAW82678	X. cumber	906	15	44.1	357	8	ADP30961	Adp30961	Human sec
834	15	44.1	212	2	AAW82685	L. cornic	907	15	44.1	361	3	ABA42153	ABA42153	Human ORF
835	15	44.1	212	2	AAW82677	X. cumber	908	15	44.1	362	8	ADN99707	ADN99707	Novel hum
836	15	44.1	213	4	AAO90922	C. glutami	909	15	44.1	362	8	ADO20269	ADO20269	Human PRO
837	15	44.1	216	7	ABO84141	Pseudomon	910	15	44.1	362	8	ABW81662	ABW81662	Tumour -a8
838	15	44.1	217	7	ABO79989	Pseudomon	911	15	44.1	366	8	ADP31670	Adp31670	Human sec
839	15	44.1	218	8	ADP69901	Rat P35NC	912	15	44.1	369	8	ADP30582	Adp30582	Human sec
840	15	44.1	222	4	ABG27806	Novel hum	913	15	44.1	377	8	ADP31659	Adp31659	Human sec
841	15	44.1	222	7	ABO77741	Pseudomon	914	15	44.1	382	8	ADP31216	Adp31216	Human sec
842	15	44.1	222	8	ADP31108	Human sec	915	15	44.1	387	4	ABW71334	ABW71334	Drosophil
843	15	44.1	223	7	ABO77463	Pseudomon	916	15	44.1	387	8	ADP31158	Adp31158	Human sec
844	15	44.1	224	8	ADP04509	Sea squit	917	15	44.1	390	8	ADP31584	Adp31584	Human sec
845	15	44.1	225	8	ADP94600	Plant ful	918	15	44.1	399	8	ADP31367	Adp31367	Human sec
846	15	44.1	226	5	AAW52385	Human CD8	919	15	44.1	399	8	ADP31317	Adp31317	Human sec
847	15	44.1	227	4	ABW58146	Drosophil	920	15	44.1	399	8	ADP31313	Adp31313	Human sec
848	15	44.1	228	8	ADP31334	Human sec	921	15	44.1	417	8	ADP30979	Adp30979	Human sec
849	15	44.1	234	8	ADP31480	Human sec	922	15	44.1	417	8	ADP31432	Adp31432	Human sec
850	15	44.1	234	8	ADP31481	Human sec	923	15	44.1	420	8	ADN20155	ADN20155	Bacterial
851	15	44.1	237	8	ADP30792	Human sec	924	15	44.1	423	8	ADP30819	Adp30819	Human sec
852	15	44.1	239	7	ABO75548	Pseudomon	925	15	44.1	423	8	ADP31323	Adp31323	Human sec
853	15	44.1	240	5	ABO75560	Breast ca	926	15	44.1	432	7	ADP31365	Adp31365	Human sec
854	15	44.1	240	6	ABR47527	Breast ca	927	15	44.1	435	7	ABO68739	ABO68739	Pseudomon
855	15	44.1	240	7	ADN38776	Cancer/an	928	15	44.1	442	8	ADP31134	Adp31134	Human sec
856	15	44.1	240	7	ABO78355	Pseudomon	929	15	44.1	443	3	AGA13372	AGA13372	Human sec
857	15	44.1	243	8	ADP30983	Human sec	930	15	44.1	456	8	ADP31224	Adp31224	Human sec
858	15	44.1	247	7	ABO72974	Pseudomon	931	15	44.1	463	7	ABO71740	ABO71740	Pseudomon
859	15	44.1	248	7	ABO71415	Pseudomon	932	15	44.1	471	8	ADP31567	Adp31567	Human sec
860	15	44.1	248	8	ADP30662	Human sec	933	15	44.1	471	8	ADP30870	Adp30870	Human sec
861	15	44.1	249	8	ADP30754	Human sec	934	15	44.1	473	8	ADP30894	Adp30894	Human sec
862	15	44.1	249	8	ADP31315	Human sec	935	15	44.1	473	8	ADP30728	Adp30728	Human sec
863	15	44.1	250	7	ABO83401	Pseudomon	936	15	44.1	478	8	ADP31157	Adp31157	Human sec
864	15	44.1	251	7	ABO73872	Pseudomon	937	15	44.1	478	8	ADP31007	Adp31007	Human sec
865	15	44.1	254	8	ADP31396	Human sec	938	15	44.1	484	8	ADP31628	Adp31628	Human sec
866	15	44.1	268	3	AAO29116	Arabidops	939	15	44.1	485	3	ABW59032	ABW59032	Breast an
867	15	44.1	268	7	ADP46627	Thalecres	940	15	44.1	491	4	ABG10165	ABG10165	Novel hum
868	15	44.1	268	8	ADP01923	Thalecres	941	15	44.1	496	4	ABW66085	ABW66085	Drosophil
869	15	44.1	268	8	ADP31697	Human sec	942	15	44.1	502	2	AAK32762	AAK32762	Thyroid N
870	15	44.1	268	8	ADP33830	Plant ful	943	15	44.1	503	8	ADP31624	Adp31624	Human sec
871	15	44.1	270	8	ADP31321	Human sec	944	15	44.1	514	8	ADP31681	Adp31681	Human sec
872	15	44.1	270	8	ADP31564	Human sec	945	15	44.1	516	8	ADP30871	Adp30871	Human sec
873	15	44.1	271	7	ABO77489	Pseudomon	946	15	44.1	520	7	ADP35609	ADP35609	Prostate
874	15	44.1	274	8	ADP31679	Human sec	947	15	44.1	520	7	ADP31845	Adp31845	Human nov
875	15	44.1	277	8	ADP30503	Human sec	948	15	44.1	520	7	ADP05027	ADP05027	Human pro
876	15	44.1	279	8	ADP31509	Human sec	949	15	44.1	522	8	ADP31081	Adp31081	Human sec
877	15	44.1	279	8	ADP30846	Human sec	950	15	44.1	522	8	ADP31018	Adp31018	Human sec
878	15	44.1	279	8	ADP30844	Human sec	951	15	44.1	524	4	AAU07370	AAU07370	G. protei
879	15	44.1	287	2	AAK73013	Functiona	952	15	44.1	525	5	ABW60120	ABW60120	Human DIT
880	15	44.1	288	8	ADP31463	Human sec	953	15	44.1	526	6	ADP31054	Adp31054	Human sec
881	15	44.1	289	2	AAK73008	Xylogluca	954	15	44.1	527	6	ABR43310	ABR43310	Human lip
882	15	44.1	293	3	AAK33374	Arabidops	955	15	44.1	534	7	ABO803340	ABO803340	Pseudomon
883	15	44.1	294	4	ABW61105	Drosophil	956	15	44.1	541	6	ABJ25731	ABJ25731	Aspergill
884	15	44.1	294	8	ADP31473	Human sec	957	15	44.1	543	8	ADP31506	Adp31506	Human sec
885	15	44.1	300	8	ADP30616	Human sec	958	15	44.1	543	8	ADP30872	Adp30872	Human sec
886	15	44.1	300	8	ADP30775	Human sec	959	15	44.1	543	8	ADP31393	Adp31393	Human sec
887	15	44.1	306	8	ADP31205	Human sec	960	15	44.1	552	8	ADP30875	Adp30875	Human sec
888	15	44.1	311	3	AAK34889	Human sec	961	15	44.1	555	8	ADP31416	Adp31416	Human sec
889	15	44.1	311	4	AAK64894	Human sec	962	15	44.1	555	8	ADP31417	Adp31417	Human sec
890	15	44.1	312	8	ADP30476	Human sec	963	15	44.1	558	8	ADP31257	Adp31257	Human sec
891	15	44.1	312	8	ADP31454	Human sec	964	15	44.1	564	6	ABJ26331	ABJ26331	Aspergill
892	15	44.1	318	7	ADP31066	Human sec	965	15	44.1	564	8	ADP31194	Adp31194	Human sec
893	15	44.1	319	7	ABO72910	Pseudomon	966	15	44.1	567	8	ADP31395	Adp31395	Human sec
894	15	44.1	321	3	AAK33373	Arabidops	967	15	44.1	569	7	ADP34357	ADP34357	Mycobacte
895	15	44.1	326	8	ADP31253	Plant ful	968	15	44.1	574	8	ADP31143	Adp31143	Rice ABC
896	15	44.1	332	7	ABO87359	Human GPC	969	15	44.1	575	8	ADP31143	Adp31143	Human sec
897	15	44.1	336	7	ABO83940	Pseudomon	970	15	44.1	576	8	ADP30522	ADP30522	Human sec
898	15	44.1	337	8	ADP31387	Human sec	971	15	44.1	583	8	ADP30553	ADP30553	Human sec
899	15	44.1	338	8	ABO58310	Human gen	972	15	44.1	585	8	ADP31392	ADP31392	Human sec
900	15	44.1	339	8	ADP30892	Human sec	973	15	44.1	585	8	ADP31445	ADP31445	Human sec

974	15	44.1	587	4	AB67608	Abb67608 Drosophil
975	15	44.1	588	8	ADP30877	Adp30877 Human sec
976	15	44.1	588	8	ADP31680	Adp31680 Human sec
977	15	44.1	588	8	ADP31699	Adp31699 Human sec
978	15	44.1	591	2	AAV15228	Aay15228 Human sec
979	15	44.1	594	8	ADP30895	Adp30895 Human sec
980	15	44.1	603	8	ADP31150	Adp31150 Human sec
981	15	44.1	603	8	ADP31645	Adp31645 Human sec
982	15	44.1	605	8	ADP30507	Adp30507 Human sec
983	15	44.1	608	8	ADP04735	Adp04735 Sea squir
984	15	44.1	609	8	ADP31263	Adp31263 Human sec
985	15	44.1	612	7	ABO69512	Abp69512 Pseudomon
986	15	44.1	612	8	ADP31064	Adp31064 Human sec
987	15	44.1	615	8	ADP31351	Adp31351 Human sec
988	15	44.1	615	8	ADP30803	Adp30803 Human sec
989	15	44.1	615	8	ADP31360	Adp31360 Human sec
990	15	44.1	617	8	ADP31657	Adp31657 Human sec
991	15	44.1	621	8	ADP30896	Adp30896 Human sec
992	15	44.1	621	8	ADP31147	Adp31147 Human sec
993	15	44.1	632	7	ABO76798	Abp76798 Pseudomon
994	15	44.1	637	8	ADP31397	Adp31397 Human sec
995	15	44.1	638	8	ADP30513	Adp30513 Human sec
996	15	44.1	640	8	ADP30977	Adp30977 Human sec
997	15	44.1	642	8	ADR86228	Adr86228 Aspergill
998	15	44.1	645	8	ADP31067	Adp31067 Human sec
999	15	44.1	648	6	ABR433306	Abt433306 Human lip
1000	15	44.1	651	8	ADP31543	Adp31543 Human sec

ALIGNMENTS

```
RESULT 1
ABO82562
ID ABO82562 strand; protein; 104 AA.
XX
AC ABO82562;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #14737.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
DR WPI; 2003-615309/58.
DR N-PSDB; ABD16133.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 31308; 455bp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
```

```
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84336 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC segdata.uspto.gov/sequence.html
SQ
Sequence 104 AA;
QY
Query Match 52.9%; Score 18; DB 7; Length 104;
Best Local Similarity 16.7%; Pred. No. 4.9e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Db 56 CSATASASTTC 67
RESULT 2
ADP30723
ID ADP30723 strand; protein; 110 AA.
XX
AC ADP30723;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1490.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410963P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
```

XX 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485245P.
PR 14-JUL-2003; 2003US-0486432P.
PR 14-JUL-2003; 2003US-0486446P.
PR 15-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486489P.
PR 15-JUL-2003; 2003US-0486560P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Hallenbeck RF, Zhang MM, Kothakota S, Haishan L, Llanemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI: 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS
XX Claim 1; SEQ ID NO 2721; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
Q Sequence 110 AA;

Query Match	Similarity	Score	DB	length
Best local	16.7%	Pred. No. 5e-05;		
Matches	2; Conservative	0; Mismatches	10;	Indels 0; Gaps 0
Oy	4 CXXXXXXXXXC	15		
Db	93 CATAAATATTAAC	104		

RESULT	3
ADP31638	
ID	ADP31638 standard; protein; 135 AA.
AC	ADP31638;
XX	
DT	12-AUG-2004 (first entry)
XX	
DE	Human secreted protein SEQ ID #2405.
XX	
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
RW	cancer; inflammatory; immune; human secreted protein.
OS	Homo sapiens.
XX	
PN	WO2004035732-A2.
XX	
PD	29-APR-2004.
XX	
PF	28-AUG-2003; 2003WO-US026760.
XX	
PR	29-AUG-2002; 2002US-0406576P.
PR	29-AUG-2002; 2002US-0406579P.
PR	29-AUG-2002; 2002US-0406585P.
PR	29-AUG-2002; 2002US-0406588P.
PR	29-AUG-2002; 2002US-0406608P.
PR	29-AUG-2002; 2002US-0406611P.
PR	29-AUG-2002; 2002US-0406612P.
PR	29-AUG-2002; 2002US-0406616P.
PR	29-AUG-2002; 2002US-0406640P.
PR	29-AUG-2002; 2002US-0406642P.
PR	29-AUG-2002; 2002US-0406646P.
PR	29-AUG-2002; 2002US-0406653P.
PR	29-AUG-2002; 2002US-0406655P.
PR	29-AUG-2002; 2002US-0406666P.
PR	29-AUG-2002; 2002US-0410946P.
PR	17-SEP-2002; 2002US-0410947P.
PR	17-SEP-2002; 2002US-0410948P.
PR	17-SEP-2002; 2002US-0410949P.
PR	17-SEP-2002; 2002US-0410953P.
PR	17-SEP-2002; 2002US-0410957P.
PR	17-SEP-2002; 2002US-0410958P.
PR	17-SEP-2002; 2002US-0410959P.
PR	17-SEP-2002; 2002US-0410960P.
PR	17-SEP-2002; 2002US-0410961P.
PR	17-SEP-2002; 2002US-0410962P.
PR	17-SEP-2002; 2002US-0411019P.
PR	17-SEP-2002; 2002US-0411022P.
PR	17-SEP-2002; 2002US-0411023P.
PR	17-SEP-2002; 2002US-0411024P.
PR	17-SEP-2002; 2002US-0411032P.
PR	17-SEP-2002; 2002US-0411035P.
PR	17-SEP-2002; 2002US-0411037P.
PR	17-SEP-2002; 2002US-0411041P.
PR	17-SEP-2002; 2002US-0411045P.
PR	17-SEP-2002; 2002US-0411046P.
PR	17-SEP-2002; 2002US-0411048P.
PR	17-SEP-2002; 2002US-0411052P.
PR	17-SEP-2002; 2002US-0411055P.
PR	17-SEP-2002; 2002US-0411073P.
PR	17-SEP-2002; 2002US-0411082P.
PR	17-SEP-2002; 2002US-0411101P.
PR	17-SEP-2002; 2002US-0411111P.
PR	18-APR-2003; 2003US-0463700P.
PR	18-APR-2003; 2003US-0463708P.
PR	18-APR-2003; 2003US-0463716P.
PR	18-APR-2003; 2003US-0463732P.
PR	02-MAY-2003; 2003US-0467199P.
PR	02-MAY-2003; 2003US-0467201P.
PR	02-MAY-2003; 2003US-0467203P.
PR	02-MAY-2003; 2003US-0467206P.
PR	19-MAY-2003; 2003US-0471306P.

```
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476009P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3636; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
XX Sequence 135 AA;
SQ
Query Match 52.9%; Score 18; DB 8; Length 135;
Best Local Similarity 16.7%; Pred. No. 5.2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXXC 15
DB 89 CATTATATAAAC 100
RESULT 4
ABO79745
ID ABO79745 standard; protein; 142 AA.
XX
XX ABO79745;
XX
XX 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polypeptide #11920.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
XX
XX US651795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX
```

```
PR 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX N-PSDB; ABD13316.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 28491; 455bp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC segdata.uspto.gov/sequence.html
XX
XX Sequence 142 AA;
SQ
Query Match 52.9%; Score 18; DB 7; Length 142;
Best Local Similarity 16.7%; Pred. No. 5.2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXXC 15
DB 88 CSASATSSATTC 99
RESULT 5
ABO73284
ID ABO73284 standard; protein; 178 AA.
XX
XX ABO73284;
XX
XX 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polypeptide #5459.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
XX
XX US651795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX
XX 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX N-PSDB; ABD06855.
XX
```

PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide.
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 22030; 455bp; English.
CC
CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-*P. aeruginosa* drugs, as templates for recombinant
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology. Sequences A067826-
CC A084396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/Sequence.html
XX
SQ Sequence 178 AA;

Query Match 52.9%; Score 18; DB 7; Length 178;
Best Local Similarity 16.7%; Pred. No. 5.5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

4 CXXXXXXXXXXC 15
92 CATATATASSTSC 103

Db

RESULT 6
ADP31527
ID ADP31527 standard; protein; 264 AA.
XX
XX ADP31527;
XX
XX 12-AUG-2004 (first entry)
XX
XX
XX Human secreted protein SEQ ID #2294.
XX
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX
XX MO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX
XX 29-AUG-2002; 2002US-0406576P.
XX
XX 29-AUG-2002; 2002US-0406579P.
XX
XX 29-AUG-2002; 2002US-0406585P.
XX
XX 29-AUG-2002; 2002US-0406588P.
XX
XX 29-AUG-2002; 2002US-0406608P.
XX
XX 29-AUG-2002; 2002US-0406611P.
XX
XX 29-AUG-2002; 2002US-0406612P.
XX
XX 29-AUG-2002; 2002US-0406616P.
XX
XX 29-AUG-2002; 2002US-0406640P.
XX
XX 29-AUG-2002; 2002US-0406642P.
XX
XX 29-AUG-2002; 2002US-0406646P.
XX
XX 29-AUG-2002; 2002US-0406653P.
XX
XX 29-AUG-2002; 2002US-0406655P.
XX
XX 29-AUG-2002; 2002US-0406666P.
XX
XX 17-SEP-2002; 2002US-0410946P.
XX
XX 17-SEP-2002; 2002US-0410947P.
XX
XX 17-SEP-2002; 2002US-0410948P.
XX
XX 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaureang PA, Behrens D,
XX Halenbeck RF, Huang MM, Kothakota S, Halsehan L, Linnemann T;
XX Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX
XX Claim 1; SEQ ID NO 3525; 428bp; English.
XX
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPWEB and is not in the specification.

```
XX      SQ      Sequence 264 AA;
Query Match          52.9%; Score 18; DB 8; Length 264;
Best Local Similarity 16.7%; Pred. No. 6e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy      4 CXXXXXXXXXXC 15
      109 CAAATATTTC 120

RESULT 7
ADP31435
ID      ADP31435 standard; protein; 270 AA.
XX
XX      ADP31435;
AC
XX      DT      12-AUG-2004 (first entry)
DE
XX      Human secreted protein SEQ ID #2202.
XX
XX      Cycostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW      cancer; inflammatory; immune; human secreted protein.
XX
XX      Homo sapiens.
OS
XX      WO2004035732-A2.
PN
XX      29-APR-2004.
PD
XX      28-AUG-2003; 2003WO-US026780.
PF
XX      29-AUG-2002; 2002US-0406576P.
PR      29-AUG-2002; 2002US-0406579P.
PR      29-AUG-2002; 2002US-0406585P.
PR      29-AUG-2002; 2002US-0406588P.
PR      29-AUG-2002; 2002US-0406608P.
PR      29-AUG-2002; 2002US-0406611P.
PR      29-AUG-2002; 2002US-0406612P.
PR      29-AUG-2002; 2002US-0406616P.
PR      29-AUG-2002; 2002US-0406640P.
PR      29-AUG-2002; 2002US-0406642P.
PR      29-AUG-2002; 2002US-0406646P.
PR      29-AUG-2002; 2002US-0406653P.
PR      29-AUG-2002; 2002US-0406655P.
PR      29-AUG-2002; 2002US-0406666P.
PR      17-SEP-2002; 2002US-0410946P.
PR      17-SEP-2002; 2002US-0410947P.
PR      17-SEP-2002; 2002US-0410948P.
PR      17-SEP-2002; 2002US-0410949P.
PR      17-SEP-2002; 2002US-0410953P.
PR      17-SEP-2002; 2002US-0410957P.
PR      17-SEP-2002; 2002US-0410958P.
PR      17-SEP-2002; 2002US-0410959P.
PR      17-SEP-2002; 2002US-0410961P.
PR      17-SEP-2002; 2002US-0410962P.
PR      17-SEP-2002; 2002US-0411019P.
PR      17-SEP-2002; 2002US-0411022P.
PR      17-SEP-2002; 2002US-0411023P.
PR      17-SEP-2002; 2002US-0411024P.
PR      17-SEP-2002; 2002US-0411032P.
PR      17-SEP-2002; 2002US-0411035P.
PR      17-SEP-2002; 2002US-0411035P.
PR      17-SEP-2002; 2002US-0411037P.
PR      17-SEP-2002; 2002US-0411041P.
PR      17-SEP-2002; 2002US-0411045P.
PR      17-SEP-2002; 2002US-0411046P.
PR      17-SEP-2002; 2002US-0411048P.
PR      17-SEP-2002; 2002US-0411052P.
PR      17-SEP-2002; 2002US-0411055P.
PR      17-SEP-2002; 2002US-0411073P.
PR      17-SEP-2002; 2002US-0411082P.
XX
```

```
PR      17-SEP-2002; 2002US-0411019P.
PR      17-SEP-2002; 2002US-0411111P.
PR      18-APR-2003; 2003US-0463708P.
PR      18-APR-2003; 2003US-0463708P.
PR      18-APR-2003; 2003US-0463716P.
PR      18-APR-2003; 2003US-0463732P.
PR      02-MAY-2003; 2003US-0467199P.
PR      02-MAY-2003; 2003US-0467201P.
PR      02-MAY-2003; 2003US-0467203P.
PR      02-MAY-2003; 2003US-0467203P.
PR      19-MAY-2003; 2003US-0471306P.
PR      19-MAY-2003; 2003US-0471336P.
PR      22-MAY-2003; 2003US-0472420P.
PR      22-MAY-2003; 2003US-0472430P.
PR      09-JUN-2003; 2003US-0476609P.
PR      09-JUN-2003; 2003US-0476641P.
PR      08-JUL-2003; 2003US-0485218P.
PR      08-JUL-2003; 2003US-0485223P.
PR      08-JUL-2003; 2003US-0485224P.
PR      08-JUL-2003; 2003US-0485325P.
PR      14-JUL-2003; 2003US-048646P.
PR      14-JUL-2003; 2003US-0486480P.
PR      15-JUL-2003; 2003US-0486891P.
PR      15-JUL-2003; 2003US-0486960P.
PR      08-AUG-2003; 2003US-0493341P.
PR      08-AUG-2003; 2003US-0493370P.
PR      08-AUG-2003; 2003US-0493573P.
PR      08-AUG-2003; 2003US-0493577P.
XX
XX      (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA
XX      Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI      Halenbeck RP, Huang MW, Kothakota S, Halsehan L, Linemann T;
PI      Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX      WPI; 2004-348438/32.
DR
XX      New nucleic acid molecule for diagnosing, preventing or treating diseases
PT      such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT      genetic, bacterial and viral diseases.
XX
XX      Claim 1; SEQ ID NO 3433; 428pp; English.
PS
XX      The present invention relates to an isolated nucleic acid molecule
CC      encoding a polypeptide which is believed to be cycostatic,
CC      antinflammatory, immunosuppressive, antibacterial and virucidal. The
CC      composition and methods are useful for diagnosing, preventing and
CC      treating diseases such as proliferative (e.g. cancer), inflammatory,
CC      immune, metabolic, genetic, bacterial and viral diseases. The present
CC      sequence represents a human secreted protein. The present sequence is
CC      available on WIPOMB and is not in the specification.
XX
XX      SQ      Sequence 270 AA;
Query Match          52.9%; Score 18; DB 8; Length 270;
Best Local Similarity 16.7%; Pred. No. 6e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy      4 CXXXXXXXXXXC 15
      179 CTATATTAAAC 190

RESULT 8
ABB65879
ID      ABB65879 standard; protein; 307 AA.
XX
XX      ABB65879;
AC
XX      DT      26-MAR-2002 (first entry)
XX
XX      Drosophila melanogaster polypeptide SEQ ID NO 24439.
DE
XX
```

KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
OS Drosophila melanogaster.
XX WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX DR N-PSDB; ABL09982.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX PS Disclosure; SEQ ID NO 24429; 21bp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-
XX CC AB872072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 307 AA;
Query Match 52.9%; Score 18; DB 4; Length 307;
Best Local Similarity 16.7%; Pred. No. 6.2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXXC 15
Db 44 CTTTTTTTTTTC 55
RESULT 9
ADP31223
ID ADP31223 standard; protein; 357 AA.
XX ADP31223;
XX AC ADP31223;
XX XX
DT 12-AUG-2004 (first entry)
XX XX
DE Human secreted protein SEQ ID #1990.
XX XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX OS Homo sapiens.
XX OS
XX PN WO2004035732-A2.
XX PN
XX PD 29-APR-2004.
XX PD
XX PF 28-AUG-2003; 2003WO-US026780.
XX PF
XX PR 29-AUG-2002; 2002US-0406576P.
XX PR 29-AUG-2002; 2002US-0406579P.
XX PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee B, Hestir K, Beaureang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Halahan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGE, Wu G, Zhang H;
XX WPI; 2004-348438/32.
DR

XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3221; 428pp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer) inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
SQ Sequence 357 AA;
Query Match 52.9%; Score 18; DB 8; Length 357;
Best Local Similarity 16.7%; Pred. No. 6.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Cy 4 CXXXXXXXXXC 15
Db 175 CTTAAATTATC 186
RESULT 10
ADP31132
ID ADP31132 standard; protein; 615 AA.
XX
AC ADP31132;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1899.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411033P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493373P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kotakota S, Hsiehan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3130; 428pp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer) inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPWEB and is not in the specification.
SQ Sequence 615 AA;
Query Match 52.9%; Score 18; DB 8; Length 615;
Best Local Similarity 16.7%; Pred. No. 7.1e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
 DB 175 CATTATTTTAC 186

RESULT 11
 AAU51580
 ID AAU51580 standard; protein; 627 AA.
 XX
 AC AAU51580;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #12476.
 XX
 KM SARPO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US012865.
 XX
 PR 21-APR-2000; 2000US-0199047P.
 XX
 PR 02-JUN-2000; 2000US-0208841P.
 XX
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'Maisonneuve J, Zhang Y, Jen S, Carter D;
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59551.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 PS Example 1; SEQ ID NO 12775; 1069pp; English.
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SARPO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 CC Sequence 627 AA;
 SQ

Query Match 52.9%; Score 18; DB 4; Length 627;
 Best Local Similarity 16.7%; Pred. No. 7.2e-05;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
 DB 46 CATTTTSATSC 57

RESULT 12
 ABM48099
 ID ABM48099 standard; protein; 627 AA.
 XX
 AC ABM48099;
 XX
 DT 20-OCT-2003 (first entry)
 XX
 DE Propionibacterium acnes predicted ORF-encoded polypeptide #12775.
 XX
 KM Acne vulgaris; antisporrhotic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO2003033515-A1.
 XX
 PD 24-APR-2003.
 XX
 PF 11-OCT-2002; 2002WO-US032727.
 XX
 PR 15-OCT-2001; 2001US-00978825.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Valliave-Douglass J;
 DR WPI; 2003-381789/36.
 DR N-PSDB; ACF64480.
 XX
 PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX
 PS Example 1; SEQ ID NO 12775; 1481pp; English.
 CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 CC Sequence 627 AA;
 SQ

```
Query Match          52.9%; Score 18; DB 6; Length 627;
Best Local Similarity 16.7%; Pred. No. 7.2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      4 CXXXXXXXXXXC 15
DB      46 CAATTTTATATSC 57

RESULT 13
ADP31142
ID ADP31142 standard; protein; 669 AA.
XX
AC ADP31142;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1909.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX
PR 29-AUG-2002; 2002US-0406579P.
XX
PR 29-AUG-2002; 2002US-0406585P.
XX
PR 29-AUG-2002; 2002US-0406588P.
XX
PR 29-AUG-2002; 2002US-0406608P.
XX
PR 29-AUG-2002; 2002US-0406611P.
XX
PR 29-AUG-2002; 2002US-0406612P.
XX
PR 29-AUG-2002; 2002US-0406616P.
XX
PR 29-AUG-2002; 2002US-0406640P.
XX
PR 29-AUG-2002; 2002US-0406642P.
XX
PR 29-AUG-2002; 2002US-0406646P.
XX
PR 29-AUG-2002; 2002US-0406653P.
XX
PR 29-AUG-2002; 2002US-0406655P.
XX
PR 29-AUG-2002; 2002US-0406666P.
XX
PR 17-SEP-2002; 2002US-0410946P.
XX
PR 17-SEP-2002; 2002US-0410947P.
XX
PR 17-SEP-2002; 2002US-0410948P.
XX
PR 17-SEP-2002; 2002US-0410949P.
XX
PR 17-SEP-2002; 2002US-0410953P.
XX
PR 17-SEP-2002; 2002US-0410957P.
XX
PR 17-SEP-2002; 2002US-0410958P.
XX
PR 17-SEP-2002; 2002US-0410959P.
XX
PR 17-SEP-2002; 2002US-0410960P.
XX
PR 17-SEP-2002; 2002US-0410961P.
XX
PR 17-SEP-2002; 2002US-0410962P.
XX
PR 17-SEP-2002; 2002US-0411019P.
XX
PR 17-SEP-2002; 2002US-0411022P.
XX
PR 17-SEP-2002; 2002US-0411023P.
XX
PR 17-SEP-2002; 2002US-0411024P.
XX
PR 17-SEP-2002; 2002US-0411029P.
XX
PR 17-SEP-2002; 2002US-0411032P.
XX
PR 17-SEP-2002; 2002US-0411035P.
XX
PR 17-SEP-2002; 2002US-0411037P.
XX
PR 17-SEP-2002; 2002US-0411041P.
XX
PR 17-SEP-2002; 2002US-0411045P.
XX
PR 17-SEP-2002; 2002US-0411046P.
XX
PR 17-SEP-2002; 2002US-0411048P.
XX
PR 17-SEP-2002; 2002US-0411052P.
XX
PR 17-SEP-2002; 2002US-0411055P.
XX
PR 17-SEP-2002; 2002US-0411057P.
XX
PR 17-SEP-2002; 2002US-0411082P.
XX
PR 17-SEP-2002; 2002US-0411010P.
XX
PR 17-SEP-2002; 2002US-0411111P.
XX
PR 18-APR-2003; 2003US-0463700P.
```

```
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 22-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 14-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Haislan L, Lammemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX PS Claim 1; SEQ ID NO 3140; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPOMEB and is not in the specification.
XX
XX SQ Sequence 669 AA;
XX
XX Query Match          52.9%; Score 18; DB 8; Length 669;
XX Best Local Similarity 16.7%; Pred. No. 7.3e-05;
XX Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
XX QY      4 CXXXXXXXXXXC 15
XX DB      652 CAATTTTAAC 663
XX
XX RESULT 14
XX ADP31215
XX ID ADP31215 standard; protein; 711 AA.
XX
XX AC ADP31215;
XX
XX DT 12-AUG-2004 (first entry)
XX
XX DE Human secreted protein SEQ ID #1982.
XX
XX KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX
```

OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-047136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486480P.
PR 14-JUL-2003; 2003US-0486486P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
CC genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3213; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
XX available on WIPOWEB and is not in the specification.
SQ Sequence 711 AA;
Query Match 52.9%; Score 18; DB 8; Length 711;
Best Local Similarity 16.7%; Pred. No. 7.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXXC 15
Db 684 CAAAAAATTC 695
RESULT 15
ID ADP31196 standard; protein; 739 AA.
XX
AC ADP31196;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1963.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411033P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485242P.
PR 08-JUL-2003; 2003US-0485252P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486691P.
PR 15-JUL-2003; 2003US-0486696P.
PR 08-AUG-2003; 2003US-049341P.
PR 08-AUG-2003; 2003US-049341P.
PR 08-AUG-2003; 2003US-0493570P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisman L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3194; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and

CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 739 AA;

Query Match 52.9%; Score 18; DB 8; Length 739;
Best Local Similarity 16.7%; Pred. No. 7.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15
DB 226 CAAAAAAAAAAC 237

RESULT 16
ADP31244
ID ADP31244 standard; protein; 771 AA.
XX
XX ADP31244;
XX
XX 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #2011.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406613P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411046P.

PR	08-JUL-2003;	2003US-0485324P.
PR	08-JUL-2003;	2003US-0485325P.
PR	14-JUL-2003;	2003US-0486446P.
PR	14-JUL-2003;	2003US-0486480P.
PR	15-JUL-2003;	2003US-0486891P.
PR	15-JUL-2003;	2003US-0486960P.
PR	08-AUG-2003;	2003US-0493341P.
PR	08-AUG-2003;	2003US-0493370P.
PR	08-AUG-2003;	2003US-0493573P.
PR	08-AUG-2003;	2003US-0493577P.
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
XX		
PI	Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;	
PI	Halenbeck RF, Huang MM, Kothakota S, Hsieh L, Linnemann T;	
PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;	
XX		
DR	WPI, 2004-348438/32.	
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	genetic, bacterial and viral diseases.	
XX		
PS	Claim 1; SEQ ID NO 3218; 428bp; English.	
XX		
CC	The present invention relates to an isolated nucleic acid molecule	
CC	encoding a polypeptide which is believed to be cytostatic.	
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	composition and methods are useful for diagnosing, preventing and	
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,	
CC	immune, metabolic, genetic, bacterial and viral diseases. The present	
CC	sequence represents a human secreted protein. The present sequence is	
CC	available on WIPOWEB and is not in the specification.	
XX		
SQ	Sequence 876 AA;	
	Query Match	52.9%; Score 18; DB 8; Length 876;
	Best Local Similarity	16.7%; Pred. No. 7.7e-05;
	Matches	2; Conservative 0; Mismatches 10; Indels 0; Gaps 0.
Cy	4 CXXXXXXXXXXC 15	
Db	28 CATATATAAATC 39	
	RESULT 18	
	ADP31688	
ID	ADP31688 standard; protein; 882 AA.	
XX		
AC	ADP31688;	
XX		
DT	12-AUG-2004 (first entry)	
XX		
DE	Human secreted protein SEQ ID #2455.	
XX		
KM	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;	
KM	cancer; Inflammatory; Immune; human secreted protein.	
XX		
OS	Homo sapiens.	
XX		
PN	WO2004035732-A2.	
XX		
PD	29-APR-2004.	
XX		
PF	28-AUG-2003; 2003WO-US026780.	
XX		
PR	29-AUG-2002; 2002US-0406576P.	
PR	29-AUG-2002; 2002US-0406579P.	
PR	29-AUG-2002; 2002US-0406585P.	
PR	29-AUG-2002; 2002US-0406588P.	
PR	29-AUG-2002; 2002US-0406608P.	
PR	29-AUG-2002; 2002US-0406611P.	
PR	29-AUG-2002; 2002US-0406612P.	

PR	29-AUG-2002	2002US-0406616P
PR	29-AUG-2002	2002US-0406649P
PR	29-AUG-2002	2002US-0406642P
PR	29-AUG-2002	2002US-0406646P
PR	29-AUG-2002	2002US-0406653P
PR	29-AUG-2002	2002US-0406655P
PR	29-AUG-2002	2002US-0406666P
PR	17-SEP-2002	2002US-0410946P
PR	17-SEP-2002	2002US-0410947P
PR	17-SEP-2002	2002US-0410948P
PR	17-SEP-2002	2002US-0410949P
PR	17-SEP-2002	2002US-0410953P
PR	17-SEP-2002	2002US-0410957P
PR	17-SEP-2002	2002US-0410958P
PR	17-SEP-2002	2002US-0410959P
PR	17-SEP-2002	2002US-0410960P
PR	17-SEP-2002	2002US-0410961P
PR	17-SEP-2002	2002US-0410962P
PR	17-SEP-2002	2002US-0411019P
PR	17-SEP-2002	2002US-0411022P
PR	17-SEP-2002	2002US-0411023P
PR	17-SEP-2002	2002US-0411024P
PR	17-SEP-2002	2002US-0411032P
PR	17-SEP-2002	2002US-0411035P
PR	17-SEP-2002	2002US-0411037P
PR	17-SEP-2002	2002US-0411041P
PR	17-SEP-2002	2002US-0411045P
PR	17-SEP-2002	2002US-0411046P
PR	17-SEP-2002	2002US-0411048P
PR	17-SEP-2002	2002US-0411052P
PR	17-SEP-2002	2002US-0411055P
PR	17-SEP-2002	2002US-0411073P
PR	17-SEP-2002	2002US-0411082P
PR	17-SEP-2002	2002US-0411101P
PR	17-SEP-2002	2002US-0411111P
PR	18-APR-2003	2003US-0463700P
PR	18-APR-2003	2003US-0463708P
PR	18-APR-2003	2003US-0463716P
PR	18-APR-2003	2003US-0463732P
PR	02-MAY-2003	2003US-0467199P
PR	02-MAY-2003	2003US-0467201P
PR	02-MAY-2003	2003US-0467203P
PR	02-MAY-2003	2003US-0467230P
PR	19-MAY-2003	2003US-0471306P
PR	19-MAY-2003	2003US-0471336P
PR	22-MAY-2003	2003US-0472420P
PR	22-MAY-2003	2003US-0472430P
PR	09-JUN-2003	2003US-0476609P
PR	09-JUN-2003	2003US-0476641P
PR	08-JUL-2003	2003US-0485218P
PR	08-JUL-2003	2003US-0485223P
PR	08-JUL-2003	2003US-0485224P
PR	08-JUL-2003	2003US-0485325P
PR	14-JUL-2003	2003US-0486446P
PR	14-JUL-2003	2003US-0486480P
PR	15-JUL-2003	2003US-0486891P
PR	15-JUL-2003	2003US-0486960P
PR	08-AUG-2003	2003US-0493341P
PR	08-AUG-2003	2003US-0493370P
PR	08-AUG-2003	2003US-0493573P
PR	08-AUG-2003	2003US-0493577P
XX		
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
XX		
PI	Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D,	
PI	Halenbeck RF, Hang MM, Kothakota S, Haisan L, Linnemann T,	
PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;	
XX		
DR	WPI; 2004-348438/32.	
XX		
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	genetic, bacterial and viral diseases.	

XX Claim 1; SEQ ID NO 3686; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.

XX
XX Sequence 882 AA;

Query Match 52.9%; Score 18; DB 8; Length 882;
Best Local Similarity 16.7%; Pred. No. 7.7e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 15
DB 763 CTTATTATTAC 774

RESULT 19
AAOI4246
ID AAO14246 standard; protein; 925 AA.

XX
XX AAO14246;

XX
XX 10-MAY-2002 (first entry)

DE Human presenilin enhancer protein pen-1B derived protein SEQ ID NO: 25.

XX Human; fruit fly; mouse; rat; cow; presenilin enhancer protein; pen;

KM Alzheimer's disease; pen-1; pen-1B; pen-2; Aph-2; amyloid beta.

XX Homo sapiens.

OS Synthetic.

XX MO200185912-A2.

XX 15-NOV-2001.

XX 03-MAY-2001; 2001WO-US014648.

XX 05-MAY-2000; 2000US-00568942.

XX (EXEL-) EXELIXIS INC.

PI Curtis DT, Francis GR, Ellis MC, Ruddy DA, Nicoll SM, McGrath GJ;

XX WPI; 2002-062245/08.

PT Presenilin enhancer proteins and polynucleotides useful for modulating
PT presenilin function and screening for an agent that modulates the
PT interaction of the protein to a binding target.

XX Disclosure; Page 67-70; 78pp; English.

XX The present invention relates to a method of detecting compounds capable
CC of altering the interaction between a presenilin enhancer protein (such
CC as pen-1, pen-1B, pen-2 and Aph-2) and presenilin. The inhibition of
CC presenilin activity causes the production of amyloid beta to be reduced
CC and thus be used in the treatment of Alzheimer's disease. The present
CC sequence is a presenilin enhancer protein described in the
CC exemplification of the invention

XX Sequence 925 AA;

Query Match 52.9%; Score 18; DB 5; Length 925;
Best Local Similarity 16.7%; Pred. No. 7.8e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 15
DB 267 CATATTATTAC 278

RESULT 20
ADP31517
ID ADP31517 standard; protein; 1044 AA.

XX ADP31517;

XX 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2284.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

XX MO2004035732-A2.

XX 29-APR-2004.

XX 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406613P.

XX 29-AUG-2002; 2002US-0406640P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406646P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

XX 29-AUG-2002; 2002US-0406666P.

XX 17-SEP-2002; 2002US-0410946P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410957P.

XX 17-SEP-2002; 2002US-0410958P.

XX 17-SEP-2002; 2002US-0410961P.

XX 17-SEP-2002; 2002US-0410962P.

XX 17-SEP-2002; 2002US-0411019P.

XX 17-SEP-2002; 2002US-0411022P.

XX 17-SEP-2002; 2002US-0411023P.

XX 17-SEP-2002; 2002US-0411024P.

XX 17-SEP-2002; 2002US-0411032P.

XX 17-SEP-2002; 2002US-0411035P.

XX 17-SEP-2002; 2002US-0411037P.

XX 17-SEP-2002; 2002US-0411041P.

XX 17-SEP-2002; 2002US-0411045P.

XX 17-SEP-2002; 2002US-0411046P.

XX 17-SEP-2002; 2002US-0411048P.

XX 17-SEP-2002; 2002US-0411052P.

XX 17-SEP-2002; 2002US-0411055P.

XX 17-SEP-2002; 2002US-0411073P.

XX 17-SEP-2002; 2002US-0411082P.

XX 17-SEP-2002; 2002US-0411101P.

XX 18-APR-2003; 2003US-0463700P.

XX 18-APR-2003; 2003US-0463708P.

XX 18-APR-2003; 2003US-0463716P.

XX 18-APR-2003; 2003US-0463732P.

XX 02-MAY-2003; 2003US-0467199P.

XX	02-MAY-2003;	2003US-0467201P.	PR
PR	02-MAY-2003;	2003US-0467203P.	PR
PR	02-MAY-2003;	2003US-0467230P.	PR
PR	19-MAY-2003;	2003US-0471306P.	PR
PR	19-MAY-2003;	2003US-0471336P.	PR
PR	22-MAY-2003;	2003US-0472420P.	PR
PR	22-MAY-2003;	2003US-0472430P.	PR
PR	09-JUN-2003;	2003US-0476609P.	PR
PR	09-JUN-2003;	2003US-0476641P.	PR
PR	08-JUL-2003;	2003US-0485218P.	PR
PR	08-JUL-2003;	2003US-0485223P.	PR
PR	08-JUL-2003;	2003US-0485224P.	PR
PR	08-JUL-2003;	2003US-0485325P.	PR
PR	14-JUL-2003;	2003US-0486446P.	PR
PR	14-JUL-2003;	2003US-0486480P.	PR
PR	15-JUL-2003;	2003US-0486891P.	PR
PR	15-JUL-2003;	2003US-0486960P.	PR
PR	08-AUG-2003;	2003US-0493341P.	PR
PR	08-AUG-2003;	2003US-0493370P.	PR
PR	08-AUG-2003;	2003US-0493573P.	PR
PR	08-AUG-2003;	2003US-0493577P.	PR
XX			
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.		
XX			
PI	Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D,		
PI	Halenbeck RF, Nhang MM, Kothakota S, Haislan L, Lannemann T;		
PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;		
XX			
DR	WPI; 2004-348438/32.		
XX			
FT	New nucleic acid molecule for diagnosing, preventing or treating diseases		
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,		
PT	genetic, bacterial and viral diseases.		
XX			
PS	Claim 1; SEQ ID NO 3515; 428bp; English.		
XX			
CC	The present invention relates to an isolated nucleic acid molecule		
CC	encoding a polypeptide which is believed to be cytostatic,		
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The		
CC	composition and methods are useful for diagnosing, preventing and		
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,		
CC	immune, metabolic, genetic, bacterial and viral diseases. The present		
CC	sequence represents a human secreted protein. The present sequence is		
CC	available on WIPOMB and is not in the specification.		
XX			
SQ	Sequence 1044 AA;		
XX			
	Query Match	52.9%; Score 18; DB 8; Length 1044;	
	Best local Similarity	16.7%; Pred. No. 8e-05;	
	Matches	2; Conservative	0; Mismatches
			10; Indels
			0; Gaps
QY			
	4 CXXXXXXXXXXC 15		
DB	325 CATATTAATTTC 336		
XX			
RESULT 21			
ADP31175			
ID	ADP31175 standard; protein; 1086 AA.		
XX			
AC	ADP31175;		
XX			
DT	12-AUG-2004 (first entry)		
XX			
DE	Human secreted protein SEQ ID #1942.		
XX			
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;		
KW	cancer; inflammatory; immune; human secreted protein.		
OS	Homo sapiens.		
XX			
PN	WO2004035732-A2.		
XX			

ED	29-APR-2004.	
XX	28-AUG-2003.	2003MO-US026780.
XX	29-AUG-2002.	2002US-0406576P.
PR	29-AUG-2002.	2002US-0406578P.
PR	29-AUG-2002.	2002US-0406585P.
PR	29-AUG-2002.	2002US-0406588P.
PR	29-AUG-2002.	2002US-0406608P.
PR	29-AUG-2002.	2002US-0406611P.
PR	29-AUG-2002.	2002US-0406612P.
PR	29-AUG-2002.	2002US-0406616P.
PR	29-AUG-2002.	2002US-0406642P.
PR	29-AUG-2002.	2002US-0406644P.
PR	29-AUG-2002.	2002US-0406646P.
PR	29-AUG-2002.	2002US-0406653P.
PR	29-AUG-2002.	2002US-0406655P.
PR	29-AUG-2002.	2002US-0406666P.
PR	29-AUG-2002.	2002US-0410956P.
PR	17-SEP-2002.	2002US-0410944P.
PR	17-SEP-2002.	2002US-0410947P.
PR	17-SEP-2002.	2002US-0410948P.
PR	17-SEP-2002.	2002US-0410949P.
PR	17-SEP-2002.	2002US-0410953P.
PR	17-SEP-2002.	2002US-0410957P.
PR	17-SEP-2002.	2002US-0411022P.
PR	17-SEP-2002.	2002US-0411022P.
PR	17-SEP-2002.	2002US-0411023P.
PR	17-SEP-2002.	2002US-0411024P.
PR	17-SEP-2002.	2002US-0411024P.
PR	17-SEP-2002.	2002US-0411032P.
PR	17-SEP-2002.	2002US-0411035P.
PR	17-SEP-2002.	2002US-0411037P.
PR	17-SEP-2002.	2002US-0411041P.
PR	17-SEP-2002.	2002US-0411045P.
PR	17-SEP-2002.	2002US-0411046P.
PR	17-SEP-2002.	2002US-0411048P.
PR	17-SEP-2002.	2002US-0411052P.
PR	17-SEP-2002.	2002US-0411055P.
PR	17-SEP-2002.	2002US-0411073P.
PR	17-SEP-2002.	2002US-0411082P.
PR	17-SEP-2002.	2002US-0411101P.
PR	17-SEP-2002.	2002US-0411111P.
PR	18-APR-2003.	2003US-0463700P.
PR	18-APR-2003.	2003US-0463708P.
PR	18-APR-2003.	2003US-0463710P.
PR	18-APR-2003.	2003US-0463716P.
PR	18-APR-2003.	2003US-0463732P.
PR	02-MAY-2003.	2003US-0467201P.
PR	02-MAY-2003.	2003US-0467201P.
PR	02-MAY-2003.	2003US-0467203P.
PR	19-MAY-2003.	2003US-0467230P.
PR	19-MAY-2003.	2003US-0471336P.
PR	19-MAY-2003.	2003US-0471336P.
PR	22-MAY-2003.	2003US-0472430P.
PR	09-JUN-2003.	2003US-0476809P.
PR	09-JUN-2003.	2003US-0476811P.
PR	08-JUL-2003.	2003US-0485232P.
PR	08-JUL-2003.	2003US-0485232P.
PR	08-JUL-2003.	2003US-0485224P.
PR	08-JUL-2003.	2003US-0485225P.
PR	14-JUL-2003.	2003US-0486454P.
PR	14-JUL-2003.	2003US-0486480P.
PR	15-JUL-2003.	2003US-0486691P.
PR	15-JUL-2003.	2003US-0486691P.
PR	08-AUG-2003.	2003US-0493341P.
PR	08-AUG-2003.	2003US-0493347P.
PR	08-AUG-2003.	2003US-0493573P.
PR	08-AUG-2003.	2003US-0493577P.
XX		

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3173; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 1086 AA;
XX
Query Match 52.9%; Score 18; DB 8; Length 1086;
Best Local Similarity 16.7%; Pred. No. 8e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXXC 15
Db 189 CATTAATTAATTC 200
RESULT 22
ADP31508
ID ADP31508 standard; protein; 1113 AA.
XX
AC ADP31508;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2275.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; Inflammatory; Immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX
PR 29-AUG-2002; 2002US-0406579P.
XX
PR 29-AUG-2002; 2002US-0406585P.
XX
PR 29-AUG-2002; 2002US-0406588P.
XX
PR 29-AUG-2002; 2002US-0406608P.
XX
PR 29-AUG-2002; 2002US-0406611P.
XX
PR 29-AUG-2002; 2002US-0406612P.
XX
PR 29-AUG-2002; 2002US-0406616P.
XX
PR 29-AUG-2002; 2002US-0406640P.
XX
PR 29-AUG-2002; 2002US-0406642P.
XX
PR 29-AUG-2002; 2002US-0406646P.
XX
PR 29-AUG-2002; 2002US-0406653P.
XX
PR 29-AUG-2002; 2002US-0406655P.
XX
PR 29-AUG-2002; 2002US-0406666P.
XX
PR 17-SEP-2002; 2002US-0410946P.
XX
PR 17-SEP-2002; 2002US-0410947P.
XX
PR 17-SEP-2002; 2002US-0410948P.
XX
PR 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3506; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.

XX SQ Sequence 1113 AA;
Query Match 52.9%; Score 18; DB 8; Length 1113;
Best Local Similarity 16.7%; Pred. No. 8.1e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXC 15
DB 18 CATATTTTAAAC 29
RESULT 23
ADA15725
ID ADA15725 standard; protein; 1128 AA.
XX ADA15725;
AC
XX
DT 06-NOV-2003 (first entry)
XX
DE C. elegans neuromuscular junction GABA receptor complex subunit #6.
XX
XX Nematode;
KW neuromuscular junction gamma-aminobutyric acid receptor complex;
KM GABA receptor; parasitic plant pathogen; agricultural industry;
KM crop protection; soil treatment.
XX Caenorhabditis elegans.
OS
XX US2003065144-A1.
PN
XX 03-APR-2003.
PD
XX 24-MAY-2002; 2002US-00156240.
PF
XX 09-NOV-1998; 9AUS-0107227P.
PR 08-NOV-1999; 99US-00436063.
XX
XX (UTAH) UNIV UTAH RES FOUND.
PA
XX Bamber BA, Jorgensen EM;
PI
XX WPI: 2003-540802/51.
DR N-PSDB; ADA15726.
XX
XX New nematode neuromuscular junction GABA receptor complex, useful for
PT crop protection or soil treatment.
XX
XX Claim 21; Page 54-56; 84pp; English.
PS
XX The present invention relates to a nematode neuromuscular junction gamma-
CC aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant
CC pathogens which can cause major damage to crops in the agricultural
CC industry. The nematode neuromuscular junction GABA receptor complex is
CC useful for crop protection or soil treatment. The present sequence
CC represents a Caenorhabditis elegans neuromuscular junction GABA receptor
CC complex subunit.
XX
XX Sequence 1128 AA;
SQ
Query Match 52.9%; Score 18; DB 6; Length 1128;
Best Local Similarity 16.7%; Pred. No. 8.1e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXC 15
DB 390 CAATATTTTTC 401
RESULT 24
ADP30741
ID ADP30741 standard; protein; 1134 AA.
XX

AC ADP30741;
XX 12-AUG-2004 (first entry)
DT
XX Human secreted protein SEQ ID #1508.
DE
XX Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX Homo sapiens.
OS
XX WO2004035732-A2.
PN
XX 29-APR-2004.
PD
XX 28-AUG-2003; 2003WO-US026780.
PF
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406615P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406665P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463715P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.

PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 2922; 428bp; English.
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
SQ Sequence 1134 AA;

Query Match 52.9%; Score 18; DB 8; Length 1134;
Best Local Similarity 16.7%; Pred. No. 8.1e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 15
DB 322 CTAATAAATAC 333

RESULT 26
ADP31046
ID ADP31046 strand; protein; 1168 AA.
AC ADP31046;
XX 12-AUG-2004 (first entry)
DT
XX
DE Human secreted protein SEQ ID #1813.
XX
XX Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; Inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX PD 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.

PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411083P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485252P.
PR 14-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 15-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kothakota S, Halsehan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3044; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPWEB and is not in the specification.
SQ Sequence 1168 AA;

Query Match 52.9%; Score 18; DB 8; Length 1168;
Best Local Similarity 16.7%; Pred. No. 8.2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 15

DB 223 CAAATTAATTAC 234

RESULT 27

ID ADP30993 standard; protein; 1191 AA.

XX ADP30993;

XX ADP30993;

XX 12-AUG-2004 (first entry)

XX Human secreted protein SEQ ID #1760.

DE Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

OS Homo sapiens.

XX WO2004035732-A2.

XX 29-APR-2004.

XX 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406640P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406646P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

XX 29-AUG-2002; 2002US-0406666P.

XX 17-SEP-2002; 2002US-0410946P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410957P.

XX 17-SEP-2002; 2002US-0410958P.

XX 17-SEP-2002; 2002US-0410959P.

XX 17-SEP-2002; 2002US-0410960P.

XX 17-SEP-2002; 2002US-0410961P.

XX 17-SEP-2002; 2002US-0410962P.

XX 17-SEP-2002; 2002US-0411019P.

XX 17-SEP-2002; 2002US-0411022P.

XX 17-SEP-2002; 2002US-0411023P.

XX 17-SEP-2002; 2002US-0411024P.

XX 17-SEP-2002; 2002US-0411032P.

XX 17-SEP-2002; 2002US-0411035P.

XX 17-SEP-2002; 2002US-0411037P.

XX 17-SEP-2002; 2002US-0411041P.

XX 17-SEP-2002; 2002US-0411045P.

XX 17-SEP-2002; 2002US-0411046P.

XX 17-SEP-2002; 2002US-0411048P.

XX 17-SEP-2002; 2002US-0411052P.

XX 17-SEP-2002; 2002US-0411055P.

XX 17-SEP-2002; 2002US-0411073P.

XX 17-SEP-2002; 2002US-0411082P.

XX 17-SEP-2002; 2002US-0411101P.

XX 17-SEP-2002; 2002US-0411111P.

XX 18-APR-2003; 2003US-0463700P.

XX 18-APR-2003; 2003US-0463708P.

XX 18-APR-2003; 2003US-0463716P.

XX 18-APR-2003; 2003US-0463732P.

XX 02-MAY-2003; 2003US-0467199P.

XX 02-MAY-2003; 2003US-0467201P.

XX 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.

PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.

PR 09-JUN-2003; 2003US-0476641P.

PR 08-JUL-2003; 2003US-0485218P.

PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485224P.

PR 08-JUL-2003; 2003US-0485252P.

PR 14-JUL-2003; 2003US-0486446P.

PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486960P.

PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.

PR 08-AUG-2003; 2003US-0493577P.

XX (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX PA

XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;

PI Halenbeck RF, Huang MM, Kotlakota S, Halsehan L, Linnemann T;

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

XX WPI; 2004-348438/32.

XX DR

XX New nucleic acid molecule for diagnosing, preventing or treating diseases

PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,

PT genetic, bacterial and viral diseases.

XX

PS Claim 1; SEQ ID NO 2991; 428bp; English.

XX

CC The present invention relates to an isolated nucleic acid molecule

CC encoding a polypeptide which is believed to be cytostatic,

CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and

CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC immune, metabolic, genetic, bacterial and viral diseases. The present

CC sequence represents a human secreted protein. The present sequence is

CC available on WIPWEB and is not in the specification.

XX

SQ Sequence 1191 AA;

Query Match 52.9%; Score 18; DB 6; Length 1191;

Best Local Similarity 16.7%; Pred. No. 8.2e-05;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 15

DB 821 CTTATTAATTATC 832

RESULT 28

ID ADP31533 standard; protein; 1260 AA.

XX ADP31533;

XX ADP31533;

XX 12-AUG-2004 (first entry)

XX Human secreted protein SEQ ID #2300.

DE Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

OS Homo sapiens.

XX WO2004035732-A2.

XX 29-APR-2004.

XX

PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-041101P.
PR 17-SEP-2002; 2002US-041111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485233P.
PR 08-JUL-2003; 2003US-0485242P.
PR 08-JUL-2003; 2003US-0485252P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486896P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493373P.
PR 08-AUG-2003; 2003US-0493377P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halebbeck RF, Huang MM, Kotnakota S, Halsehan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3531; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic.
CC antinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 1260 AA;
XX
Query Match 52.9%; Score 18; DB 8; Length 1260;
Best Local Similarity 16.7%; Pred. No. 8.3e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
OY 4 CXXXXXXXXXC 15
Db 757 CAATTAATAC 768
RESULT 29
ADP30675
ID ADP30675 standard; protein; 1289 AA.
ADP30675;
AC ADP30675;
XX
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1442.
XX
XX
KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-047136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halebek RF, Huang MM, Kochakota S, Halsehan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX MPI; 2004-348438/32.
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1, SEQ ID NO 2673; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPWEB and is not in the specification.
SQ Sequence 1289 AA;

Query Match 52.9%; Score 18; DB 8; Length 1289;
Best Local Similarity 16.7%; Pred. No. 8.3e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXXC 15
Db 1076 CATATATATATTC 1087
RESULT 30
ADP31357
ID ADP31357 standard; protein; 1437 AA.
XX
AC ADP31357;
XX
XX 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2124.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; Inflammatory; Immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
PD 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411101P.
XX 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493373P.
PR 08-AUG-2003; 2003US-0493577P.
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3355; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic.
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 1437 AA;
XX
Query Match 52.9%; Score 18; DB 8; Length 1437;
Best Local Similarity 16.7%; Pred. No. 8.5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXXC 15
DB 206 CAATTAATATATC 217
XX
RESULT 31
ADP31177
ID ADP31177 standard; protein; 1454 AA.
XX
AC ADP31177;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1944.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
cancer; inflammatory; immune; human secreted protein.

XX OS Homo sapiens.
XX PN
XX WO2004035732-A2.
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.

PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
PS
PS Claim 1; SEQ ID NO 3175; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 1454 AA;
XX
Query Match 52.9%; Score 18; DB 8; Length 1454;
Best Local Similarity 16.7%; Pred. No. 8.5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXXC 15
Db 1094 CATTATATATAC 1105
XX
RESULT 32
ADP30557
ID ADP30557 standard; protein; 1480 AA.
XX
AC ADP30557;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1324.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493373P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D;
XX Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;
XX Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2555; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.

XX
SQ Sequence 1480 AA;

Query Match 52.9%; Score 18; DB 8; Length 1480;
Best Local Similarity 16.7%; Pred. No. 8.6e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

Db 1286 CTAAATATTAAC 1297

RESULT 33

ID ABB09437 standard; protein; 1588 AA.

XX ABB09437;

DT 01-JUL-2002 (first entry)

DE H. influenzae DXR related polypeptide sequence.

XX DXR; reductoisomerase; enzyme; non-mevalonate isoprenoid; menaquinone;

KW ubiquinone; virulence; ear infection; conjunctivitis; meningitis;
KM pneumonia; conjunctivitis; bacteraemia; sinusitis; pleural empyema;
KM endocarditis; epiglottitis.

XX Haemophilus influenzae.

OS
FH Key Location/Qualifiers

FT Region 241..1431

FT /note="region that appears to be accidentally inserted
FT into the sequence, consisting the DXR encoding DNA
FT sequence represented as an amino acid sequence in three
FT letter code"

PD WO200211673-A2.

XX 14-FEB-2002.

PF 09-AUG-2001; 2001WO-US024950.

XX 09-AUG-2000; 2000US-0223909P.

PA (SMTK) SMITHKLINE BEECHAM CORP.

PA (SMTK) SMITHKLINE BEECHAM PLC.

PI Jaworski DD, Payne DJ, Slater-Radosci CE, Yan K;

DR WPI; 2002-241698/29.

PT Modulating Haemophilus influenzae DXR reductoisomerase enzyme activity,
PT useful for treating mammals or tissues infected with H. influenzae (e.g.
PT ear infections or pneumonia) by contacting the enzyme with a modulator of
PT its activity.

PS Disclosure; Page 40-44; 44pp; English.

XX The invention relates to modulating an activity of a DXR reductoisomerase
XX enzyme of Haemophilus influenzae, comprising contacting the enzyme with a
XX compound that modulates non-mevalonate isoprenoid biosynthesis -
XX synthesis of menaquinone or ubiquinone. Compounds of the invention act as
XX virucides. The method is useful for treating a mammal or mammalian tissue
XX infected with H. influenzae having DXR reductoisomerase enzyme, e.g. a
XX human or a domestic animal. In particular, the method is useful for
XX treating ear infections, conjunctivitis, meningitis, pneumonia,
XX conjunctivitis, bacteraemia, sinusitis, pleural empyema, endocarditis and
XX epiglottitis. The current sequence represents a H. influenzae DXR

CC reductoisomerase enzyme related polypeptide sequence. Note: The current
CC sequence contains within it the amino acid sequence given in record
CC ABB09436 (DXR enzyme), but this is broken up by a large insertion that
CC appears to be accidentally inserted into the sequence, consisting the DXR
CC encoding DNA sequence represented as an amino acid sequence in three
CC letter code

XX
SQ Sequence 1588 AA;

Query Match 52.9%; Score 18; DB 5; Length 1588;
Best Local Similarity 16.7%; Pred. No. 8.7e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

Db 431 CTAAATTTTAC 442

RESULT 34

ID ADP30660 standard; protein; 1617 AA.

XX ADP30660;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #1427.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

XX WO2004035732-A2.

XX 29-APR-2004.

PD 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406640P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406646P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

XX 17-SEP-2002; 2002US-0410946P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410957P.

XX 17-SEP-2002; 2002US-0410958P.

XX 17-SEP-2002; 2002US-0410959P.

XX 17-SEP-2002; 2002US-0410960P.

XX 17-SEP-2002; 2002US-0410961P.

XX 17-SEP-2002; 2002US-0410962P.

XX 17-SEP-2002; 2002US-0411019P.

XX 17-SEP-2002; 2002US-0411022P.

XX 17-SEP-2002; 2002US-0411023P.

XX 17-SEP-2002; 2002US-0411024P.

XX 17-SEP-2002; 2002US-0411032P.

XX 17-SEP-2002; 2002US-0411035P.

XX 17-SEP-2002; 2002US-0411037P.

XX 17-SEP-2002; 2002US-0411041P.

XX 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.
 PR 17-SEP-2002; 2002US-0411048P.
 PR 17-SEP-2002; 2002US-0411052P.
 PR 17-SEP-2002; 2002US-0411055P.
 PR 17-SEP-2002; 2002US-0411073P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411101P.
 PR 17-SEP-2002; 2002US-0411111P.
 PR 18-APR-2003; 2003US-0463700P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 19-MAY-2003; 2003US-0471336P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 09-JUN-2003; 2003US-0476609P.
 PR 09-JUN-2003; 2003US-0476641P.
 PR 08-JUL-2003; 2003US-0485218P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 14-JUL-2003; 2003US-0486448P.
 PR 15-JUL-2003; 2003US-0486891P.
 PR 15-JUL-2003; 2003US-0486960P.
 PR 08-AUG-2003; 2003US-0493341P.
 PR 08-AUG-2003; 2003US-0493370P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.
 XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
 PA William LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
 PI Halenbeck RF, Huang MM, Kothakota S, Haispan L, Linnemann T;
 PI Pierce K, Wang Y, Wong JGF, Wu G, Zhang H;
 XX WPI; 2004-348438/32.
 DR
 XX
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 PT genetic, bacterial and viral diseases.
 PS Claim 1; SEQ ID NO 2658; 428bp; English.
 XX
 CC The present invention relates to an isolated nucleic acid molecule
 CC encoding a polypeptide which is believed to be cytostatic,
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is
 CC available on WIPOMEB and is not in the specification.
 XX
 SQ Sequence 1617 AA;
 Query Match 52.9%; Score 18; DB 8; Length 1617;
 Best Local Similarity 16.7%; Pred. No. 8.7e-05;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 4 CXXXXXXXXXXC 15
 DB 71 CTATATTAATTC 82
 RESULT 35
 ADA15715
 ID ADA15715 standard; protein; 1652 AA.
 XX

AC ADA15715;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE C. elegans neuromuscular junction GABA receptor complex subunit #1.
 XX
 KW Nematode;
 KW neuromuscular junction gamma-aminobutyric acid receptor complex;
 KW GABA receptor; parasitic plant pathogen; agricultural industry;
 KW crop protection; soil treatment.
 XX
 OS Caenorhabditis elegans.
 XX
 PN US2003065144-A1.
 XX
 PD 03-APR-2003.
 XX
 PF 24-MAY-2002; 2002US-00156240.
 XX
 PR 09-NOV-1998; 98US-0107727P.
 PR 08-NOV-1999; 99US-00436063.
 XX
 PA (UTAH) UNITV UTAH RES FOUND.
 XX
 PI Bamber BA, Jorgensen EM;
 XX
 DR WPI; 2003-540802/51.
 XX
 PT New nematode neuromuscular junction GABA receptor complex, useful for
 PT crop protection or soil treatment.
 PS Claim 21; Page 20-24; 84bp; English.
 XX
 CC The present invention relates to a nematode neuromuscular junction gamma-
 CC aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant
 CC pathogens which can cause major damage to crops in the agricultural
 CC industry. The nematode neuromuscular junction GABA receptor complex is
 CC useful for crop protection or soil treatment. The present sequence
 CC represents a Caenorhabditis elegans neuromuscular junction GABA receptor
 CC complex subunit.
 XX
 SQ Sequence 1652 AA;
 Query Match 52.9%; Score 18; DB 6; Length 1652;
 Best Local Similarity 16.7%; Pred. No. 8.8e-05;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 4 CXXXXXXXXXXC 15
 DB 891 CAATATTTTTC 902
 RESULT 36
 ADP30654
 ID ADP30654 standard; protein; 1725 AA.
 XX
 AC ADP30654;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human secreted protein SEQ ID #1421.
 XX
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
 KW cancer; inflammatory; immune; human secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO2004035732-A2.
 XX
 PD 29-APR-2004.
 XX
 PF 28-AUG-2003; 2003WO-US026780.
 XX

PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471366P.
PR 19-MAY-2003; 2003US-0471367P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485252P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2652; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
SQ Sequence 1725 AA:
Query Match 52.9%; Score 18; DB 8; Length 1725;
Best Local Similarity 16.7%; Pred. No. 8.9e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXXC 15
DB 828 CTAATTTTAAAC 839
RESULT 37
ADP30642
ID ADP30642 standard; protein; 1833 AA.
XX
XX ADP30642;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1409.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
XX WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0405588P.
XX 29-AUG-2002; 2002US-0406088P.
XX 29-AUG-2002; 2002US-0406112P.
XX 29-AUG-2002; 2002US-0406116P.
XX 29-AUG-2002; 2002US-0406402P.
XX 29-AUG-2002; 2002US-0406442P.
XX 29-AUG-2002; 2002US-0406446P.
XX 29-AUG-2002; 2002US-0406453P.
XX 29-AUG-2002; 2002US-0406555P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411011P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-047136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kotchakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2640; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
XX Sequence 1833 AA;
SQ

Query Match

52.9%; Score 18; DB 8; Length 1833;

Best Local Similarity 16.7%; Pred. No. 9e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXC 15
Db 462 CTAATTTAAAC 473
RESULT 38
ADP31178
ID ADP31178 standard; protein; 2088 AA.
XX
XX ADP31178;
AC XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1945.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
PD XX
PF 28-AUG-2003; 2003WC-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406589P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411012P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411101P.
XX 17-SEP-2002; 2002US-0411111P.
XX 18-APR-2003; 2003US-0463700P.
XX 18-APR-2003; 2003US-0463708P.

PR	18-APR-2003;	2003US-0463716P.
PR	18-APR-2003;	2003US--0467323P.
PR	02-MAY-2003;	2003US-0467195P.
PR	02-MAY-2003;	2003US-0467201P.
PR	02-MAY-2003;	2003US-0467203P.
PR	02-MAY-2003;	2003US-0467230P.
PR	19-MAY-2003;	2003US-0471306P.
PR	19-MAY-2003;	2003US-0471336P.
PR	22-MAY-2003;	2003US-0472420P.
PR	22-MAY-2003;	2003US-0472430P.
PR	09-JUN-2003;	2003US-0476609P.
PR	09-JUN-2003;	2003US-0476641P.
PR	08-JUL-2003;	2003US-0485218P.
PR	08-JUL-2003;	2003US-0485223P.
PR	08-JUL-2003;	2003US-0485325P.
PR	14-JUL-2003;	2003US-0486446P.
PR	14-JUL-2003;	2003US-0486480P.
PR	15-JUL-2003;	2003US-0486891P.
PR	15-JUL-2003;	2003US-0486960P.
PR	08-AUG-2003;	2003US-0493341P.
PR	08-AUG-2003;	2003US-0493370P.
PR	08-AUG-2003;	2003US-0493372P.
PR	08-AUG-2003;	2003US-0493577P.
XX	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
XX		
PI	Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;	
PI	Halenbeck RF, Huang MM, Kothakota S, Halshan L, Linnemann T;	
PI	Pierce K, Wang Y, Wong JGF, Wu G, Zhang H;	
XX		
DR	WPT; 2004-348438/32.	
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	genetic, bacterial and viral diseases.	
XX		
PS	Claim 1; SEQ ID NO 3176; 428bp; English.	
CC	The present invention relates to an isolated nucleic acid molecule	
CC	encoding a polypeptide which is believed to be cytostatic,	
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	composition and methods are useful for diagnosing, preventing and	
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,	
CC	immune, metabolic, genetic, bacterial and viral diseases. The present	
CC	sequence represents a human secreted protein. The present sequence is	
XX	available on WIPOWEB and is not in the specification.	
SQ	Sequence 2088 AA;	
Query Match	52.9%;	Score 18; DB 8; Length 2088;
Best Local Similarity	16.7%;	Pred. No. 9.2e-05;
Matches	2; Conservative	0; Mismatches 10; Indels 0; Gaps 0;
CY	4 CXXXXXXXXXC 15	
Dd	1094 CATTTATATTAC 1105	
RESULT 39		
ADP31327		
ID	ADP31327 standard; protein; 2127 AA.	
XX	ADP31327;	
AC		
XX	12-AUG-2004 (first entry)	
DT		
XX		
DE	Human secreted protein SEQ ID #2094.	
XX		
KM	Cytostatic; AntiInflammatory; Immunosuppressive; Antibacterial; Virucide;	
KW	cancer; Inflammation; Immune; human secreted protein.	
OS	Homo sapiens.	

XX	PN	WO2004.035732-A2.	
XX	PD	29-APR-2004.	
XX	PF	28-AUG-2003.	2003WO-US026780.
XX	XX		
PR	29-AUG-2002.	2002US-0406576P.	
PR	29-AUG-2002.	2002US-0406579P.	
PR	29-AUG-2002.	2002US-0406585P.	
PR	29-AUG-2002.	2002US-0406588P.	
PR	29-AUG-2002.	2002US-0406589P.	
PR	29-AUG-2002.	2002US-0406611P.	
PR	29-AUG-2002.	2002US-0406612P.	
PR	29-AUG-2002.	2002US-0406615P.	
PR	29-AUG-2002.	2002US-0406640P.	
PR	29-AUG-2002.	2002US-0406642P.	
PR	29-AUG-2002.	2002US-0406646P.	
PR	29-AUG-2002.	2002US-0406653P.	
PR	29-AUG-2002.	2002US-0406655P.	
PR	29-AUG-2002.	2002US-0406666P.	
PR	17-SEP-2002.	2002US-0410946P.	
PR	17-SEP-2002.	2002US-0410947P.	
PR	17-SEP-2002.	2002US-0410948P.	
PR	17-SEP-2002.	2002US-0410949P.	
PR	17-SEP-2002.	2002US-0410950P.	
PR	17-SEP-2002.	2002US-0410951P.	
PR	17-SEP-2002.	2002US-0410952P.	
PR	17-SEP-2002.	2002US-0410953P.	
PR	17-SEP-2002.	2002US-0410954P.	
PR	17-SEP-2002.	2002US-0410955P.	
PR	17-SEP-2002.	2002US-0410956P.	
PR	17-SEP-2002.	2002US-0410957P.	
PR	17-SEP-2002.	2002US-0410958P.	
PR	17-SEP-2002.	2002US-0410959P.	
PR	17-SEP-2002.	2002US-0411032P.	
PR	17-SEP-2002.	2002US-0411033P.	
PR	17-SEP-2002.	2002US-0411035P.	
PR	17-SEP-2002.	2002US-0411037P.	
PR	17-SEP-2002.	2002US-0411041P.	
PR	17-SEP-2002.	2002US-0411045P.	
PR	17-SEP-2002.	2002US-0411046P.	
PR	17-SEP-2002.	2002US-0411048P.	
PR	17-SEP-2002.	2002US-0411052P.	
PR	17-SEP-2002.	2002US-0411053P.	
PR	17-SEP-2002.	2002US-0411055P.	
PR	17-SEP-2002.	2002US-0411073P.	
PR	17-SEP-2002.	2002US-0411082P.	
PR	17-SEP-2002.	2002US-0411111P.	
PR	17-SEP-2002.	2002US-0411113P.	
PR	18-APR-2003.	2003US-0463700P.	
PR	18-APR-2003.	2003US-0463708P.	
PR	18-APR-2003.	2003US-0463716P.	
PR	18-APR-2003.	2003US-0463732P.	
PR	02-MAY-2003.	2003US-0467192P.	
PR	02-MAY-2003.	2003US-0467201P.	
PR	02-MAY-2003.	2003US-0467203P.	
PR	02-MAY-2003.	2003US-0467230P.	
PR	19-MAY-2003.	2003US-0471306P.	
PR	19-MAY-2003.	2003US-0471336P.	
PR	22-MAY-2003.	2003US-0472420P.	
PR	22-MAY-2003.	2003US-0472430P.	
PR	09-JUN-2003.	2003US-0476609P.	
PR	09-JUN-2003.	2003US-0476641P.	
PR	08-JUL-2003.	2003US-0485218P.	
PR	08-JUL-2003.	2003US-0485223P.	
PR	08-JUL-2003.	2003US-0485224P.	
PR	08-JUL-2003.	2003US-0485335P.	
PR	14-JUL-2003.	2003US-0486446P.	
PR	14-JUL-2003.	2003US-0486480P.	
PR	15-JUL-2003.	2003US-0486891P.	
PR	15-JUL-2003.	2003US-0486960P.	
PR	08-AUG-2003.	2003US-0493341P.	
PR	08-AUG-2003.	2003US-0493370P.	

CC aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant
CC pathogens which can cause major damage to crops in the agricultural
CC industry. The nematode neuromuscular junction GABA receptor complex is
CC useful for crop protection or soil treatment. The present sequence
CC represents a Caenorhabditis elegans neuromuscular junction GABA receptor
CC complex subunit.

XX
SQ Sequence 2508 AA;

Query Match 52.9%; Score 18; DB 6; Length 2508;
Best Local Similarity 16.7%; Pred. No. 9.6e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 1775 CAAATATTTTTC 1786

RESULT 42

ID ADA15717 standard; protein; 2544 AA.

XX ADA15717;

AC 06-NOV-2003 (first entry)

XX C. elegans neuromuscular junction GABA receptor complex subunit #2.

XX Nematode;

KW neuromuscular junction gamma-aminobutyric acid receptor complex;

KM GABA receptor; parasitic plant pathogen; agricultural industry;

KW crop protection; soil treatment.

OS Caenorhabditis elegans.

XX US2003065144-A1.

XX 03-APR-2003.

XX 24-MAY-2002; 2002US-00156240.

XX 09-NOV-1998; 98US-0107727P.

XX 08-NOV-1999; 99US-00436063.

XX (UTAH) UNIV UTAH RES FOUND.

XX Bamber BA, Jorgensen EM;

XX WPI; 2003-540802/51.

XX N-PSDB; ADA15718.

XX New nematode neuromuscular junction GABA receptor complex, useful for

XX crop protection or soil treatment.

XX Claim 21; Page 25-31; 84pp; English.

XX The present invention relates to a nematode neuromuscular junction gamma-

XX aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant

XX pathogens which can cause major damage to crops in the agricultural

XX industry. The nematode neuromuscular junction GABA receptor complex is

XX useful for crop protection or soil treatment. The present sequence

XX represents a Caenorhabditis elegans neuromuscular junction GABA receptor

XX complex subunit.

XX Sequence 2544 AA;

RESULT 43
ADA15723
ID ADA15723 standard; protein; 2601 AA.

XX ADA15723;

AC 06-NOV-2003 (first entry)

XX C. elegans neuromuscular junction GABA receptor complex subunit #5.

XX Nematode;

KW neuromuscular junction gamma-aminobutyric acid receptor complex;

KM GABA receptor; parasitic plant pathogen; agricultural industry;

KW crop protection; soil treatment.

OS Caenorhabditis elegans.

XX US2003065144-A1.

XX 03-APR-2003.

XX 24-MAY-2002; 2002US-00156240.

XX 09-NOV-1998; 98US-0107727P.

XX 08-NOV-1999; 99US-00436063.

XX (UTAH) UNIV UTAH RES FOUND.

XX Bamber BA, Jorgensen EM;

XX WPI; 2003-540802/51.

XX N-PSDB; ADA15724.

XX New nematode neuromuscular junction GABA receptor complex, useful for

XX crop protection or soil treatment.

XX Claim 21; Page 46-52; 84pp; English.

XX The present invention relates to a nematode neuromuscular junction gamma-

XX aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant

XX pathogens which can cause major damage to crops in the agricultural

XX industry. The nematode neuromuscular junction GABA receptor complex is

XX useful for crop protection or soil treatment. The present sequence

XX represents a Caenorhabditis elegans neuromuscular junction GABA receptor

XX complex subunit.

XX Sequence 2601 AA;

Query Match 52.9%; Score 18; DB 6; Length 2601;
Best Local Similarity 16.7%; Pred. No. 9.6e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 1251 CAAATAATATTC 1262

RESULT 44
ADP31299
ID ADP31299 standard; protein; 2833 AA.
XX ADP31299;
XX 12-AUG-2004 (first entry)
XX Human secreted protein SEQ ID #2066.
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX Homo sapiens.

XX	W02004035732-A2.	
XX	29-APR-2004.	
XX	28-AUG-2003;	2003W0-US026780
XX	29-AUG-2002;	2002US-0406576P
XX	29-AUG-2002;	2002US-0406579P
XX	29-AUG-2002;	2002US-0406585P
XX	29-AUG-2002;	2002US-0406588P
XX	29-AUG-2002;	2002US-0406608P
XX	29-AUG-2002;	2002US-0406611P
XX	29-AUG-2002;	2002US-0406616P
XX	29-AUG-2002;	2002US-0406642P
XX	29-AUG-2002;	2002US-0406646P
XX	29-AUG-2002;	2002US-0406653P
XX	29-AUG-2002;	2002US-0406655P
XX	29-AUG-2002;	2002US-0406666P
XX	29-AUG-2002;	2002US-0406966P
XX	17-SEP-2002;	2002US-0410947P
XX	17-SEP-2002;	2002US-0410948P
XX	17-SEP-2002;	2002US-0410949P
XX	17-SEP-2002;	2002US-0410953P
XX	17-SEP-2002;	2002US-0410957P
XX	17-SEP-2002;	2002US-0410958P
XX	17-SEP-2002;	2002US-0410959P
XX	17-SEP-2002;	2002US-0410960P
XX	17-SEP-2002;	2002US-0410962P
XX	17-SEP-2002;	2002US-0410962P
XX	17-SEP-2002;	2002US-0411019P
XX	17-SEP-2002;	2002US-0411022P
XX	17-SEP-2002;	2002US-0411023P
XX	17-SEP-2002;	2002US-0411024P
XX	17-SEP-2002;	2002US-0411032P
XX	17-SEP-2002;	2002US-0411035P
XX	17-SEP-2002;	2002US-0411037P
XX	17-SEP-2002;	2002US-0411041P
XX	17-SEP-2002;	2002US-0411045P
XX	17-SEP-2002;	2002US-0411046P
XX	17-SEP-2002;	2002US-0411048P
XX	17-SEP-2002;	2002US-0411052P
XX	17-SEP-2002;	2002US-0411055P
XX	17-SEP-2002;	2002US-0411073P
XX	17-SEP-2002;	2002US-0411082P
XX	17-SEP-2002;	2002US-0411101P
XX	17-SEP-2002;	2002US-0411111P
XX	18-APR-2003;	2003US-0463700P
XX	18-APR-2003;	2003US-0463708P
XX	18-APR-2003;	2003US-0463716P
XX	18-APR-2003;	2003US-0463732P
XX	02-MAY-2003;	2003US-0467135P
XX	02-MAY-2003;	2003US-0467201P
XX	02-MAY-2003;	2003US-0467203P
XX	02-MAY-2003;	2003US-0467230P
XX	09-JUN-2003;	2003US-0476609P
XX	09-JUN-2003;	2003US-0476619P
XX	08-JUL-2003;	2003US-0485218P
XX	08-JUL-2003;	2003US-0485223P
XX	08-JUL-2003;	2003US-0485224P
XX	08-JUL-2003;	2003US-0485335P
XX	14-JUL-2003;	2003US-0486446P
XX	14-JUL-2003;	2003US-0486480P
XX	15-JUL-2003;	2003US-0486891P
XX	15-JUL-2003;	2003US-0486960P
XX	08-AUG-2003;	2003US-0493331P
XX	08-AUG-2003;	2003US-0493370P

PR		08-AUG-2003;	2003US-0493573P.
PR		08-AUG-2003;	2003US-0493577P.
XX			
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.		
XX			
PI	Williams LT, Chu K, Lee B, Healtir K, Beaurang PA, Behrens D;		
PI	Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linemann T;		
PI	Plececk K, Wang Y, Wong JGF, Wu G, Zhang H;		
DR	WPI; 2004-348438/32.		
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases		
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,		
PT	genetic, bacterial and viral diseases.		
PS	Claim 1; SEQ ID NO 3297; 428bp; English.		
XX			
CC	The present invention relates to an isolated nucleic acid molecule		
CC	encoding a polypeptide which is believed to be cytostatic,		
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The		
CC	composition and methods are useful for diagnosing, preventing and		
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,		
CC	immune, metabolic, genetic, bacterial and viral diseases. The present		
CC	sequence represents a human secreted protein. The present sequence is		
CC	avaliable on WIPOMEB and is not in the specification.		
SQ	Sequence 2833 AA;		
	Query Match	52.9%; Score 18; DB 8; Length 2833;	
	Best Local Similarity	16.7%; Pred. NO. 9.8e-05;	
	Matches	2; Conservative	0; Mismatches
			10; Indels
			0; Gaps
OY	4 CXXXXXXXXXKC 15		
DB	1876 CATAAATTATAC 1887		
	RESULT 45		
	ADP30572		
ID	ADP30572 standard; protein; 2835 AA.		
XX			
AC	ADP30572;		
XX			
DT	12-AUG-2004 (first entry)		
XX			
DE	Human secreted protein SEQ ID #1339.		
XX			
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;		
KW	cancer; inflammatory; immune; human secreted protein.		
XX			
OS	Homo sapiens.		
XX			
FN	WO2004035732-A2.		
XX			
PD	29-APR-2004.		
XX			
PE	28-AUG-2003; 2003WO-US026780.		
XX			
XX			
PR	29-AUG-2002; 2002US-0406576P.		
PR	29-AUG-2002; 2002US-0406578P.		
PR	29-AUG-2002; 2002US-0406585P.		
PR	29-AUG-2002; 2002US-0406588P.		
PR	29-AUG-2002; 2002US-0406608P.		
PR	29-AUG-2002; 2002US-0406611P.		
PR	29-AUG-2002; 2002US-0406612P.		
PR	29-AUG-2002; 2002US-0406616P.		
PR	29-AUG-2002; 2002US-0406640P.		
PR	29-AUG-2002; 2002US-0406642P.		
PR	29-AUG-2002; 2002US-0406646P.		
PR	29-AUG-2002; 2002US-0406653P.		
PR	29-AUG-2002; 2002US-0406655P.		
PR	29-AUG-2002; 2002US-0406666P.		
PR	17-SEP-2002; 2002US-0410946P.		

PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halesbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 2570; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
XX
SQ Sequence 2835 AA;
Query Match 52.9%; Score 18; DB 8; Length 2835;
Best Local Similarity 16.7%; Pred. No. 9.8e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXXC 15
Db 884 CAAATATATTC 895
RESULT 46
ADP30667
ID ADP30667 standard; protein; 3411 AA.
XX
XX ADP30667;
AC
XX
XX 12-AUG-2004 (first entry)
DT XX
XX
DE Human secreted protein SEQ ID #1434.
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
XX PN MO2004035732-A2.
PD
XX 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493373P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halebek R, Huang MM, Kochakota S, Haisan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

PS Claim 1: SEQ ID NO 2665; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.

XX Sequence 3411 AA;

Query Match 52.9%; Score 18; DB 8; Length 3411;
Best Local Similarity 16.7%; Pred. No. 0.0001;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 2826 CAATTTTATTTT 2837

RESULT 47

ABU88256
ID ABU88256 standard; protein; 4440 AA.
XX ABU88256;
AC
XX
DT 07-JUL-2003 (first entry)

XX Novel human secreted and transmembrane pro polypeptide #4.
DE Human; secreted and transmembrane protein. PRO; gene therapy;
XX tumour necrosis factor-alpha release; TNF-alpha release;
XX chondrocyte proliferation; chondrocyte differentiation; tumour;
XX adrenal tumour; lung tumour; colon tumour; breast tumour;
XX prostate tumour; rectal tumour; cervical tumour; liver tumour.
OS Homo sapiens.
XX
XX US2003032127-A1.
XX
XX 13-FEB-2003.
PD
XX
PF 26-JUN-2002; 2002US-00183012.
XX
XX 18-SEP-1997; 97US-0059263P.
XX 18-SEP-1997; 97US-0059266P.
XX 17-OCT-1997; 97US-0062250P.
XX 21-OCT-1997; 97US-0063486P.
XX 24-OCT-1997; 97US-0063120P.
XX 24-OCT-1997; 97US-0063121P.
XX 28-OCT-1997; 97US-0063540P.
XX 28-OCT-1997; 97US-0063541P.
XX 28-OCT-1997; 97US-0063544P.
XX 29-OCT-1997; 97US-0063734P.
XX 31-OCT-1997; 97US-0063870P.
XX 31-OCT-1997; 97US-0064103P.
XX 13-NOV-1997; 97US-0065119P.
XX 21-NOV-1997; 97US-0065120P.
XX 24-NOV-1997; 97US-0066466P.
XX 24-NOV-1997; 97US-0066772P.
XX 11-DEC-1997; 97US-0069335P.
XX 12-DEC-1997; 97US-0069425P.
XX 17-DEC-1997; 97US-0069870P.
XX 18-DEC-1997; 97US-0068017P.
XX 18-MAR-1998; 98US-0077450P.
XX 11-MAR-1998; 98US-0077632P.
XX 11-MAR-1998; 98US-0077649P.
XX 20-MAR-1998; 98US-0078886P.
XX 20-MAR-1998; 98US-0078939P.
XX 27-MAR-1998; 98US-0079664P.
XX 27-MAR-1998; 98US-0079786P.
XX 31-MAR-1998; 98US-0080107P.
XX 31-MAR-1998; 98US-0080194P.
XX 01-APR-1998; 98US-0080327P.
XX 01-APR-1998; 98US-0080333P.
XX 08-APR-1998; 98US-0081049P.
XX 08-APR-1998; 98US-0081070P.
XX 09-APR-1998; 98US-0081195P.
XX 15-APR-1998; 98US-0082568P.
XX 21-APR-1998; 98US-0082569P.
XX 21-APR-1998; 98US-0082704P.
XX 22-APR-1998; 98US-0083797P.
XX 28-APR-1998; 98US-0083722P.
XX 29-APR-1998; 98US-0083495P.
XX 29-APR-1998; 98US-0083496P.
XX 29-APR-1998; 98US-0083499P.
XX 29-APR-1998; 98US-0083559P.
XX 05-MAY-1998; 98US-0084366P.
XX 06-MAY-1998; 98US-0084414P.
XX 07-MAY-1998; 98US-0084639P.
XX 07-MAY-1998; 98US-0084640P.
XX 07-MAY-1998; 98US-0084643P.
XX 15-MAY-1998; 98US-0085579P.
XX 15-MAY-1998; 98US-0085580P.
XX 15-MAY-1998; 98US-0085582P.
XX 15-MAY-1998; 98US-0085700P.
XX 18-MAY-1998; 98US-0086023P.
XX 22-MAY-1998; 98US-0086392P.

PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095398P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.

PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-01019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.

Query Match 52.9%; Score 18; DB 6; Length 4440;
Best Local Similarity 16.7%; Pred. No. 0.00011;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXC 15
DB 3858 CATTTTAATAC 3869

RESULT 48
ABU90135
ID ABU90135 standard; protein; 4440 AA.
AC ABU90135;
XX 11-AUG-2003 (first entry)
DB Novel human secreted and transmembrane PRO protein #4.
XX Human; gene therapy; tissue typing; tumour; chondrocyte proliferation;
KW chondrocyte differentiation; tumour necrosis factor-alpha release;
XX affinity purification.

OS Homo sapiens.
XX
XX US2003036147-A1.
XX
XX 20-FEB-2003.
XX
XX 02-JUL-2002; 2002US-00187741.
XX
PR 18-SRP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069670P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088555P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P*.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095598P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096157P.
PR 17-AUG-1998; 98US-0096476P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.

PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0098602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-01019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102865P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98US-00168978.

Query Match 52.9%; Score 18; DB 6; Length 4440;
Best Local Similarity 16.7%; Pred. No. 0.00011;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXC 15
DB 3858 CATTTTTATAC 3869

RESULT 49
ABU96437
ID ABU96437 standard; protein; 4440 AA.

XX ABU96437;

XX 25-JUL-2003 (first entry)

XX Novel human secreted and transmembrane protein #4.

XX Human; secreted and transmembrane protein; PRO; transgenic animal;

KW knockout; chromosome identification; tissue typing; tumour;

KW chondrocyte proliferation; chondrocyte differentiation;

XX tumor necrosis factor-alpha release stimulator.

OS Homo sapiens.

XX US2003036144-A1.
XX 20-FEB-2003.
XX

PF 01-JUL-2002; 2002US-00187601..
XX 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 28-OCT-1997; 97US-0063734P.
PR 29-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 31-OCT-1997; 97US-0065311P.
PR 13-NOV-1997; 97US-0066120P.
PR 21-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 24-NOV-1997; 97US-0069335P.
PR 11-DEC-1997; 97US-0069425P.
PR 12-DEC-1997; 97US-0069870P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0077450P.
PR 10-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082568P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 06-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086233P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.

PR 09-JUN-1998; 98US-0088655P.
 PR 10-JUN-1998; 98US-0088722P.
 PR 10-JUN-1998; 98US-0088738P.
 PR 10-JUN-1998; 98US-0088740P.
 PR 10-JUN-1998; 98US-0088811P.
 PR 10-JUN-1998; 98US-0088824P.
 PR 10-JUN-1998; 98US-0088825P.
 PR 10-JUN-1998; 98US-0088826P.
 PR 11-JUN-1998; 98US-0088861P.
 PR 11-JUN-1998; 98US-0088863P.
 PR 11-JUN-1998; 98US-0088876P.
 PR 12-JUN-1998; 98US-0089000P.
 PR 12-JUN-1998; 98US-0089105P.
 PR 16-JUN-1998; 98US-0089512P.
 PR 16-JUN-1998; 98US-0089514P.
 PR 17-JUN-1998; 98US-0089538P.
 PR 17-JUN-1998; 98US-0089588P.
 PR 17-JUN-1998; 98US-0089653P.
 PR 18-JUN-1998; 98US-0089908P.
 PR 19-JUN-1998; 98US-0089922P.
 PR 22-JUN-1998; 98US-0090246P.
 PR 22-JUN-1998; 98US-0090252P.
 PR 22-JUN-1998; 98US-0090254P.
 PR 24-JUN-1998; 98US-0090429P.
 PR 24-JUN-1998; 98US-0090435P.
 PR 24-JUN-1998; 98US-0090444P.
 PR 24-JUN-1998; 98US-0090461P.
 PR 24-JUN-1998; 98US-0090555P.
 PR 24-JUN-1998; 98US-0090540P.
 PR 25-JUN-1998; 98US-0090676P.
 PR 25-JUN-1998; 98US-0090687P.
 PR 25-JUN-1998; 98US-0090688P.
 PR 25-JUN-1998; 98US-0090690P.
 PR 25-JUN-1998; 98US-0090694P.
 PR 25-JUN-1998; 98US-0090695P.
 PR 25-JUN-1998; 98US-0090696P.
 PR 26-JUN-1998; 98US-00105413.
 PR 26-JUN-1998; 98US-0090862P.
 PR 26-JUN-1998; 98US-0090863P.
 PR 26-JUN-1998; 98US-0091010P.
 PR 01-JUL-1998; 98US-0091359P.
 PR 01-JUL-1998; 98US-0091544P.
 PR 02-JUL-1998; 98US-0091478P.
 PR 02-JUL-1998; 98US-0091486P.
 PR 02-JUL-1998; 98US-0091626P.
 PR 02-JUL-1998; 98US-0091628P.
 PR 02-JUL-1998; 98US-0091632P.
 PR 04-JUL-1998; 98US-0094006P.
 PR 04-AUG-1998; 98US-0095282P.
 PR 10-AUG-1998; 98US-0095998P.
 PR 10-AUG-1998; 98US-0096012P.
 PR 17-AUG-1998; 98US-0096757P.
 PR 17-AUG-1998; 98US-0096766P.
 PR 17-AUG-1998; 98US-0096867P.
 PR 17-AUG-1998; 98US-0096891P.
 PR 17-AUG-1998; 98US-0096897P.
 PR 18-AUG-1998; 98US-0096849P.
 PR 18-AUG-1998; 98US-0096859P.
 PR 18-AUG-1998; 98US-0097022P.
 PR 26-AUG-1998; 98US-0097952P.
 PR 26-AUG-1998; 98US-0097954P.
 PR 26-AUG-1998; 98US-0097955P.
 PR 26-AUG-1998; 98US-0097971P.
 PR 26-AUG-1998; 98US-0097974P.
 PR 26-AUG-1998; 98US-0098014P.
 PR 01-SEP-1998; 98US-0098716P.
 PR 01-SEP-1998; 98US-0098723P.
 PR 02-SEP-1998; 98US-0098803P.
 PR 02-SEP-1998; 98US-0098821P.
 PR 02-SEP-1998; 98US-0098843P.
 PR 09-SEP-1998; 98US-0099602P.
 PR 10-SEP-1998; 98US-0099741P.
 PR 10-SEP-1998; 98US-0099754P.

PR 10-SEP-1998; 98US-0099763P.
 PR 10-SEP-1998; 98US-0099812P.
 PR 15-SEP-1998; 98US-0100388P.
 PR 16-SEP-1998; 98US-0100662P.
 PR 16-SEP-1998; 98US-0100664P.
 PR 16-SEP-1998; 98US-0101751P.
 PR 16-SEP-1998; 98US-0101751P.
 PR 17-SEP-1998; 98US-0100683P.
 PR 17-SEP-1998; 98US-0100684P.
 PR 17-SEP-1998; 98US-0100919P.
 PR 17-SEP-1998; 98US-0100930P.
 PR 18-SEP-1998; 98US-0100849P.
 PR 18-SEP-1998; 98US-0101014P.
 PR 18-SEP-1998; 98US-0101068P.
 PR 23-SEP-1998; 98US-0101471P.
 PR 23-SEP-1998; 98US-0101472P.
 PR 23-SEP-1998; 98US-0101475P.
 PR 23-SEP-1998; 98US-0101477P.
 PR 24-SEP-1998; 98US-0101738P.
 PR 24-SEP-1998; 98US-0101739P.
 PR 24-SEP-1998; 98US-0101743P.
 PR 24-SEP-1998; 98US-0101922P.
 PR 25-SEP-1998; 98US-0101786P.
 PR 25-SEP-1998; 98US-0102207P.
 PR 29-SEP-1998; 98US-0102207P.
 PR 29-SEP-1998; 98US-0102330P.
 PR 29-SEP-1998; 98US-0102331P.
 PR 30-SEP-1998; 98US-0102487P.
 PR 30-SEP-1998; 98US-0102570P.
 PR 30-SEP-1998; 98US-0102571P.
 PR 01-OCT-1998; 98US-0102684P.
 PR 01-OCT-1998; 98US-0102687P.
 PR 02-OCT-1998; 98US-0102958P.
 PR 06-OCT-1998; 98US-0103258P.
 PR 06-OCT-1998; 98US-0103449P.

Query Match 52.9%; Score 18; DB 6; Length 440;
 Best Local Similarity 16.7%; Pred. No. 0.00011;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

Db 3858 CATTTTTAATAC 3869

RESULT 50

ABU99046
 ID ABU99046 standard; protein; 4440 AA.

XX
 AC ABU99046;

XX
 DT 01-AUG-2003 (first entry)

XX
 DE Novel human secreted and transmembrane protein #4.

XX
 KW Human; secreted and transmembrane protein; PRO; cytosolic; gene therapy;
 KW chondrocyte stimulator; tumour; adrenal tumour; lung tumour;
 KW colon tumour; breast tumour; prostate tumour; rectal tumour;
 KW cervical tumour; liver tumour; TNF-alpha release;
 KW tumour necrosis factor alpha release; chondrocyte cell proliferation;
 KW chondrocyte cell differentiation; pharmaceutical; diagnostic; biosensor;
 KW bioreactor.

XX
 OS Homo sapiens.

XX
 PN US2003013153-A1.

XX
 PD 16-JAN-2003.

XX
 PF 19-JUN-2002; 2002US-00175737.

XX
 XX 18-SEP-1997; 97US-0059263P.

PR 18-SEP-1997; 97US-0059266P.

PR 17-OCT-1997; 97US-0062350P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082977P.
PR 28-APR-1998; 98US-0083122P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084314P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.

PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0088909P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 18-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090687P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 02-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095988P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 26-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.

PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WC-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100649P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match

52.9%; Score 18; DB 6; Length 4440;

Best Local Similarity 16.7%; Pred. No. 0.00011; Mismatches 10; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 15
Db 3858 CATTTTAATAC 3869

Search completed: January 4, 2006, 15:56:12
Job time : 132.183 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 ; Search time 18.6261 Seconds
(without alignments)
92.983 Million cell updates/sec

Title: US-09-932-322-5
Perfect score: 34
Sequence: 1 XXXXXXXXXXXXXXXXXX 18

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	52.9	307	1	salivary glue prot
2	17	50.0	245	2	hypothetical prote
3	16	47.1	152	2	hypothetical prote
4	16	47.1	244	2	hypothetical prote
5	16	47.1	248	2	hypothetical prote
6	16	47.1	257	2	hypothetical prote
7	16	47.1	539	2	glutathione syntha
8	16	47.1	667	2	trichozole surfac
9	16	47.1	770	2	related to C2H2 zi
10	16	47.1	1041	2	chitin synthase (E
11	15	44.1	154	2	probable prepilin
12	15	44.1	226	2	cytochrome c553i p
13	15	44.1	248	2	hypothetical prote
14	15	44.1	268	2	probable NAM (no a
15	15	44.1	273	2	probable finger pr
16	15	44.1	283	2	conserved hypothet
17	15	44.1	287	2	xyloglucan endo-1,
18	15	44.1	289	2	CG10208 protein (A
19	15	44.1	318	2	hypothetical prote
20	15	44.1	357	2	Sd protein - fruit
21	15	44.1	363	2	hypothetical prote
22	15	44.1	392	2	sucrose cleavage p
23	15	44.1	409	2	gamma-glutamyl pho
24	15	44.1	420	2	unknown protein 7
25	15	44.1	423	2	hypothetical prote
26	15	44.1	503	2	pyruvate (or indol
27	15	44.1	569	2	metalloprotease
28	15	44.1	727	2	box A-binding fac
29	15	44.1	779	1	

30	15	44.1	1003	1	PYVZAM	spheroidin precurs
31	15	44.1	1043	2	T19734	hypothetical prote
32	15	44.1	1186	2	T18210	delta endotoxin -
33	15	44.1	1364	2	T10236	xanthine dehydrog
34	15	44.1	1451	2	S65571	patern formation
35	15	44.1	2098	2	T25888	hypothetical prote
36	15	44.1	2176	2	T13806	coucan gene protei
37	15	44.1	2187	2	S60224	polyketide synth
38	15	44.1	2533	2	T28675	alpha-51D immobili
39	15	44.1	2533	2	T28674	alpha-51D-immobili
40	15	44.1	2543	2	T21687	surface antigen - P
41	15	44.1	2703	1	A24420	notch protein - fr
42	15	44.1	4767	2	T31345	hypothetical prote
43	15	44.1	7829	2	T15789	hypothetical prote
44	15	44.1	7962	2	I18346	elastic titin - hu
45	14	41.2	21	2	D22565	R-phycocerythrin be
46	14	41.2	88	2	G82754	hypothetical prote
47	14	41.2	92	2	I54781	fibroblast growth
48	14	41.2	97	2	T28947	hypothetical prote
49	14	41.2	100	2	C72775	hypothetical prote
50	14	41.2	103	2	T14780	hypothetical prote
51	14	41.2	107	2	S73440	hypothetical prote
52	14	41.2	107	2	S46078	probable membrane
53	14	41.2	107	4	S57388	hypothetical prote
54	14	41.2	108	2	S67638	probable membrane
55	14	41.2	129	2	S48641	lysozyme (EC 3.2.1
56	14	41.2	130	2	S57896	polyketide synthas
57	14	41.2	131	2	A43980	neurophysin 2 [val
58	14	41.2	133	2	S44539	probable membrane
59	14	41.2	135	2	T15610	hypothetical prote
60	14	41.2	143	2	A54793	pointcullin precurs
61	14	41.2	154	2	E72513	hypothetical prote
62	14	41.2	156	2	T18339	hypothetical prote
63	14	41.2	156	2	AH0865	prepilin peptidase
64	14	41.2	156	2	C65065	prepilin peptidase
65	14	41.2	158	2	S35201	serine proteinase
66	14	41.2	159	2	B82692	hypothetical prote
67	14	41.2	164	2	T28741	hypothetical prote
68	14	41.2	165	2	T28742	hypothetical prote
69	14	41.2	166	2	T33970	hypothetical prote
70	14	41.2	177	2	S73288	R-phycocerythrin be
71	14	41.2	177	2	S30932	hypothetical prote
72	14	41.2	177	2	F72706	hypothetical prote
73	14	41.2	195	2	S39967	hypothetical prote
74	14	41.2	197	2	T28739	nitrilase homolog
75	14	41.2	199	2	S50353	hypothetical prote
76	14	41.2	204	2	T41853	AcMNPV orf115 - Bo
77	14	41.2	205	2	T10384	hypothetical prote
78	14	41.2	209	2	JC7239	peroxiredoxin V -
79	14	41.2	214	2	C70812	probable lipQ prot
80	14	41.2	225	2	T22417	hypothetical prote
81	14	41.2	232	2	A60095	larval glue protei
82	14	41.2	249	2	T31837	hypothetical prote
83	14	41.2	250	2	T31836	hypothetical prote
84	14	41.2	250	2	T31835	hypothetical prote
85	14	41.2	259	2	H71353	conserved hypothet
86	14	41.2	264	2	C70603	hypothetical prote
87	14	41.2	264	2	S42793	pds-1 protein - Ca
88	14	41.2	268	2	T05971	probable receptor -
89	14	41.2	270	2	T27786	hypothetical prote
90	14	41.2	274	2	D82784	phosphomethylprim
91	14	41.2	275	2	T05969	receptor-like kina
92	14	41.2	276	2	AE0860	hypothetical prote
93	14	41.2	286	1	C42053	gap junction prote
94	14	41.2	288	2	T21158	hypothetical prote
95	14	41.2	293	2	T31840	hypothetical prote
96	14	41.2	298	2	S15524	variant surface gl
97	14	41.2	300	2	T05729	hypothetical prote
98	14	41.2	300	2	S16344	variant surface gl
99	14	41.2	312	2	T31834	hypothetical prote
100	14	41.2	315	2	T34528	hypothetical prote
101	14	41.2	328	2	T16065	iron-sulfur cofact
102	14	41.2	329	2	T15823	hypothetical prote

103	14	41.2	337	2	147079	176	14	41.2	706	2	T49700	related to Ap-1-1i
104	14	41.2	343	1	HMITVA	177	14	41.2	710	2	T21339	hypothetical prote
105	14	41.2	343	2	S45321	178	14	41.2	712	2	T18195	gag protein - slk
106	14	41.2	343	2	U02370	179	14	41.2	722	2	T00049	hemocyte protein A
107	14	41.2	343	2	U02372	180	14	41.2	729	2	S36605	nitrate reductase
108	14	41.2	343	2	U02371	181	14	41.2	742	2	T25415	hypothetical prote
109	14	41.2	344	1	A27701	182	14	41.2	758	2	A96687	hypothetical prote
110	14	41.2	344	1	A32141	183	14	41.2	769	2	A41029	integrin beta-8 ch
111	14	41.2	344	2	I45894	184	14	41.2	776	2	T02584	probable protein k
112	14	41.2	344	2	I57698	185	14	41.2	780	2	T03156	ribonucleoside-dip
113	14	41.2	344	2	P00319	186	14	41.2	791	2	B46171	coat protein - lei
114	14	41.2	344	2	P00321	187	14	41.2	794	2	G84688	hypothetical prote
115	14	41.2	345	2	A05279	188	14	41.2	856	2	D86258	protein F501.10 l
116	14	41.2	347	2	S12916	189	14	41.2	880	2	AE0179	probable ATPase ch
117	14	41.2	349	2	AD3650	190	14	41.2	882	2	AE0119	Cip ATPase limport
118	14	41.2	353	2	S16785	191	14	41.2	927	2	T21772	hypothetical prote
119	14	41.2	353	2	S65086	192	14	41.2	946	2	C86549	polymorphic outer
120	14	41.2	353	2	UC4875	193	14	41.2	946	2	D81594	polymorphic membra
121	14	41.2	354	2	A35788	194	14	41.2	946	2	C72075	polymorphic outer
122	14	41.2	357	2	S58168	195	14	41.2	988	2	T51054	related to alpha-a
123	14	41.2	358	2	S69886	196	14	41.2	989	2	T48845	insulin II gene en
124	14	41.2	364	2	S57937	197	14	41.2	1011	2	S75806	hemolysin secretio
125	14	41.2	370	2	S32173	198	14	41.2	1088	2	A69493	cysteine proteinas
126	14	41.2	378	2	R66307	199	14	41.2	1105	2	T18295	Ap-3 adaptor compl
127	14	41.2	384	2	D86598	200	14	41.2	1165	2	S58236	pyruvate (flavodox
128	14	41.2	384	2	D71478	201	14	41.2	1206	2	E86445	hypothetical prote
129	14	41.2	384	2	C81741	202	14	41.2	1229	2	S14199	probable adenylate
130	14	41.2	384	2	P72025	203	14	41.2	1230	2	E64664	outer membrane pro
131	14	41.2	385	2	T25452	204	14	41.2	1237	2	D71850	probable outer mem
132	14	41.2	390	2	E88925	205	14	41.2	1237	2	A34598	ecdysone-induced p
133	14	41.2	392	2	T46418	206	14	41.2	1268	2	G85154	hypothetical prote
134	14	41.2	424	2	S74788	207	14	41.2	1287	2	T42658	hypothetical prote
135	14	41.2	425	2	T48724	208	14	41.2	1313	2	T30548	adenylate cyclase
136	14	41.2	433	2	T50395	209	14	41.2	1331	2	T18310	receptor-adenylate
137	14	41.2	439	2	A36385	210	14	41.2	1332	2	T15670	hypothetical prote
138	14	41.2	465	1	UC1318	211	14	41.2	1344	2	AD2103	two-component hydr
139	14	41.2	469	2	C84783	212	14	41.2	1357	2	S61187	probable membrane
140	14	41.2	474	2	T27297	213	14	41.2	1361	2	C71403	hypothetical 15.1k
141	14	41.2	478	2	T24805	214	14	41.2	1367	1	S48478	glucan 1,4-alpha-g
142	14	41.2	478	2	T41408	215	14	41.2	1380	2	T18309	receptor-adenylate
143	14	41.2	479	2	S65466	216	14	41.2	1387	2	A96771	hypothetical prote
144	14	41.2	494	2	T51529	217	14	41.2	1394	2	B34598	ecdysone-induced p
145	14	41.2	495	1	S46284	218	14	41.2	1429	2	T13720	gene expanded prot
146	14	41.2	496	1	U02152	219	14	41.2	1443	2	S05979	steroid hormone re
147	14	41.2	499	2	S22571	220	14	41.2	1450	2	T30273	hypothetical prote
148	14	41.2	501	2	G85097	221	14	41.2	1797	2	T21889	hypothetical prote
149	14	41.2	515	2	A25048	222	14	41.2	1805	2	T21888	hypothetical prote
150	14	41.2	527	2	A98180	223	14	41.2	1820	2	S71853	genome polyprotein
151	14	41.2	527	2	AC3107	224	14	41.2	1823	2	S28974	vitellogenin precu
152	14	41.2	529	2	A44508	225	14	41.2	1986	2	S28353	probable polyketid
153	14	41.2	537	2	H88087	226	14	41.2	2109	2	T17490	polyketide synthas
154	14	41.2	559	1	A29941	227	14	41.2	2257	2	A57710	acetyl-CoA carboxy
155	14	41.2	559	1	A35029	228	14	41.2	2311	2	T06161	acetyl-CoA carboxy
156	14	41.2	561	2	UC2447	229	14	41.2	2325	2	T02235	acetyl-CoA carboxy
157	14	41.2	563	1	S49889	230	14	41.2	2395	1	S50820	surface protein ty
158	14	41.2	566	1	HMITV	231	14	41.2	2437	2	S42612	transmembrane prot
159	14	41.2	566	1	HMITVR	232	14	41.2	2458	2	T17420	probable polyketid
160	14	41.2	566	1	HMITVR	233	14	41.2	2531	2	T16743	hypothetical prote
161	14	41.2	566	2	S69887	234	14	41.2	2704	2	S09118	G surface protein
162	14	41.2	566	2	S69888	235	14	41.2	2718	2	A23475	G surface protein
163	14	41.2	567	1	B45865	236	14	41.2	3133	2	S52093	hemocytin - silkw
164	14	41.2	575	1	VCLJHD	237	14	41.2	3164	1	WMBEH6	hemocytin - silkw
165	14	41.2	593	2	T10301	238	14	41.2	3172	2	S22012	erythronolide synt
166	14	41.2	595	2	A42086	239	14	41.2	3178	2	S13595	6-deoxyerythronoli
167	14	41.2	598	2	T46327	240	14	41.2	3229	2	S27852	probable cell-surf
168	14	41.2	606	2	S35427	241	14	41.2	3345	2	T13423	hypothetical prote
169	14	41.2	606	2	D86232	242	14	41.2	3491	2	T43231	probable 6-deoxyer
170	14	41.2	620	2	T50232	243	14	41.2	4427	2	PN0637	polyketide synthas
171	14	41.2	649	2	C90113	244	13	38.2	39	2	B31940	microtubule-associ
172	14	41.2	653	2	S40962	245	13	38.2	55	1	MTNHB2	neurotoxin B-II -
173	14	41.2	661	2	D85361	246	13	38.2	57	2	B38180	hypothetical prote
174	14	41.2	686	2	A59348	247	13	38.2	61	1	T2NU2Y	short toxin CW-2 -
175	14	41.2	697	1	TVPPTL	248	13	38.2	63	1	XTBO1	acrosin inhibitor

249	13	38.2	66	2	A28644	kappa-neurotoxin -	322	13	38.2	201	2	T16181	hypothetical prote
250	13	38.2	66	2	JN0408	neurotoxin Os-1 -	323	13	38.2	202	2	C90742	hypothetical prote
251	13	38.2	71	2	T00008	copy number contro	324	13	38.2	202	2	P85592	hypothetical prote
252	13	38.2	79	2	T27605	hypothetical prote	325	13	38.2	204	2	D72864	AcOrf-115 protein
253	13	38.2	81	2	E42465	probable transcrip	326	13	38.2	205	2	H71457	hypothetical prote
254	13	38.2	98	2	E84599	hypothetical prote	327	13	38.2	208	2	C97654	hypothetical prote
255	13	38.2	101	2	T42262	hypothetical prote	328	13	38.2	208	2	A12877	conserved hypotet
256	13	38.2	102	2	AE1357	hypothetical prote	329	13	38.2	209	2	A82680	2-amino-4-hydroxy-
257	13	38.2	102	2	AE1727	hypothetical prote	330	13	38.2	210	2	S75973	hypothetical prote
258	13	38.2	107	2	T27988	hypothetical prote	331	13	38.2	213	2	S7585	uracase accessory p
259	13	38.2	108	2	B95314	hypothetical prote	332	13	38.2	216	2	T50615	hypothetical prote
260	13	38.2	111	2	AC0733	probable exported	333	13	38.2	216	2	C90162	acetyltransferase,
261	13	38.2	111	2	T20452	hypothetical prote	334	13	38.2	217	2	A86600	CT734 hypothetical
262	13	38.2	111	2	T34565	hypothetical prote	335	13	38.2	217	2	A72024	conserved hypotet
263	13	38.2	112	2	T51142	nonspecific lipid-	336	13	38.2	217	2	A42644	neuB protein - Bac
264	13	38.2	112	2	T48770	hypothetical prote	337	13	38.2	217	2	T48642	hypothetical prote
265	13	38.2	114	2	S57692	probable membrane	338	13	38.2	217	2	UC7875	L-rhamnose-binding
266	13	38.2	114	2	F69264	hypothetical prote	339	13	38.2	220	2	T25789	hypothetical prote
267	13	38.2	114	2	T01930	ethylene-response	340	13	38.2	222	2	T27939	hypothetical prote
268	13	38.2	117	2	T02169	homeobox protein H	341	13	38.2	225	2	S45356	probable serine pr
269	13	38.2	119	2	B45937	early chorion prot	342	13	38.2	226	2	G95287	probable [imported
270	13	38.2	119	2	S24292	chorion protein -	343	13	38.2	226	2	T21380	hypothetical prote
271	13	38.2	119	2	S24294	chorion protein -	344	13	38.2	227	2	A33937	Ig light chain (13
272	13	38.2	119	2	S24291	chorion protein -	345	13	38.2	227	2	A55472	ornithine decarbox
273	13	38.2	121	2	S24293	chorion protein -	346	13	38.2	229	1	MMVYPM	25k protein - pota
274	13	38.2	121	2	S30170	mercury resistance	347	13	38.2	233	2	T21496	hypothetical prote
275	13	38.2	125	2	AB0071	probable exported	348	13	38.2	235	2	T16916	hypothetical prote
276	13	38.2	127	2	S27915	nucleic acid-bindi	349	13	38.2	236	2	R84307	glucose 1-dehydrog
277	13	38.2	127	2	H90091	hypothetical prote	350	13	38.2	237	2	T25152	hypothetical prote
278	13	38.2	129	2	T47958	hypothetical prote	351	13	38.2	238	2	S77659	inner cell wall ma
279	13	38.2	132	1	MF1V2T	matrix protein M2	352	13	38.2	238	2	T22098	hypothetical prote
280	13	38.2	141	2	T33983	hypothetical prote	353	13	38.2	241	2	T24782	hypothetical prote
281	13	38.2	144	2	A36324	growth arrest-spec	354	13	38.2	242	2	A96794	unknown protein F1
282	13	38.2	145	2	H87436	hypothetical prote	355	13	38.2	242	2	H65136	hypothetical prote
283	13	38.2	145	2	T33037	hypothetical prote	356	13	38.2	243	2	F86006	hypothetical prote
284	13	38.2	146	1	W9BPT3	gene 19.2 protein	357	13	38.2	243	2	G91160	hypothetical prote
285	13	38.2	148	2	S19462	probable membrane	358	13	38.2	246	2	T00628	hypothetical prote
286	13	38.2	150	2	S66926	hypothetical prote	359	13	38.2	246	2	T25532	hypothetical prote
287	13	38.2	155	2	C62890	conserved hypotet	360	13	38.2	247	2	T21406	hypothetical prote
288	13	38.2	156	2	C91089	prepillin peptidase	361	13	38.2	247	2	S72859	hypothetical prote
289	13	38.2	156	2	D75258	hypothetical prote	362	13	38.2	248	2	S05007	calcium channel al
290	13	38.2	156	2	B85934	prepillin peptidase	363	13	38.2	248	2	T11913	hypothetical prote
291	13	38.2	157	2	JN0503	hypothetical 17K p	364	13	38.2	249	2	T21785	hypothetical prote
292	13	38.2	160	2	JN0503	peripherall myelin	365	13	38.2	249	2	T27782	hypothetical prote
293	13	38.2	160	2	S21721	growth arrest-spec	366	13	38.2	250	2	T36131	hypothetical prote
294	13	38.2	160	2	A41144	growth arrest-rele	367	13	38.2	250	2	T26725	hypothetical prote
295	13	38.2	162	2	T11978	phycocyanin alpha	368	13	38.2	252	2	T29527	hypothetical prote
296	13	38.2	163	2	D63883	hypothetical prote	369	13	38.2	254	2	E83619	probable transpor
297	13	38.2	163	2	B44813	probable molybdenu	370	13	38.2	255	2	T03408	MDS box protein -
298	13	38.2	166	1	NVBO2	vaopressin / neur	371	13	38.2	255	2	T34163	hypothetical prote
299	13	38.2	166	2	F81696	hypothetical prote	372	13	38.2	257	2	D71544	hypothetical prote
300	13	38.2	166	2	T32701	hypothetical prote	373	13	38.2	259	2	T27783	hypothetical prote
301	13	38.2	168	1	NVRT2	vaopressin / neur	374	13	38.2	261	2	S70006	finger protein zfo
302	13	38.2	168	2	B43755	vaopressin / neur	375	13	38.2	262	2	S01914	El membrane glycop
303	13	38.2	168	2	T18944	melanoma ubiqutin	376	13	38.2	262	2	G86938	conserved hypotet
304	13	38.2	171	2	H84923	hypothetical prote	377	13	38.2	265	2	T33059	hypothetical prote
305	13	38.2	173	2	F96543	hypothetical prote	378	13	38.2	266	2	T10816	throxine deiodina
306	13	38.2	173	2	A71144	hypothetical prote	379	13	38.2	268	2	T43486	hypothetical prote
307	13	38.2	176	2	T30961	hypothetical prote	380	13	38.2	269	2	T21407	hypothetical prote
308	13	38.2	180	2	A95283	lipuin peroxidase	381	13	38.2	271	2	E86352	protein T26F17.16
309	13	38.2	180	2	PS0011	lipuin peroxidase	382	13	38.2	272	2	T19418	hypothetical prote
310	13	38.2	180	2	E71337	hypothetical prote	383	13	38.2	274	2	T39087	hypothetical prote
311	13	38.2	182	2	S29611	pollen-preferentia	384	13	38.2	275	2	A34866	T-cell surface pro
312	13	38.2	183	2	T18923	microfibril-associ	385	13	38.2	275	2	S08464	T-cell allergen
313	13	38.2	183	2	A54151	microfibril-associ	386	13	38.2	275	2	T05822	hypothetical prote
314	13	38.2	183	2	A49313	microfibril-associ	387	13	38.2	276	2	T43541	purine nucleotide
315	13	38.2	185	2	S57419	PAR-1a protein - c	388	13	38.2	276	2	AD1987	hypothetical prote
316	13	38.2	185	2	S62698	photoassimilate-re	389	13	38.2	277	2	H97266	mind family ATPase
317	13	38.2	185	2	S62699	photoassimilate-re	390	13	38.2	278	2	T16494	hypothetical prote
318	13	38.2	198	2	JF0316	proctoan RNase, Dd	391	13	38.2	280	2	A42424	chitinase (BC 3.2.
319	13	38.2	199	2	T48099	hypothetical prote	392	13	38.2	287	1	A45168	probable 3-oxoacyl
320	13	38.2	199	2	T32267	hypothetical prote	393	13	38.2	287	2	T04236	xyloglucan endo-1,
321	13	38.2	201	2	S37847	hypothetical prote	394	13	38.2	289	2	AE2006	hypothetical prote

395	13	38.2	291	2	T28846	hypothetical prote	468	13	38.2	380	1	D65201	glycerol dehydroge
396	13	38.2	294	1	G65106	lipoprotein precu	469	13	38.2	380	2	B91238	glycerol dehydroge
397	13	38.2	294	2	AC0901	conserved hypotnet	470	13	38.2	380	2	F86085	glycerol dehydroge
398	13	38.2	294	2	D91134	lipoprotein precu	471	13	38.2	382	2	B86268	FljB4.1 protein -
399	13	38.2	294	2	G65979	lipoprotein precu	472	13	38.2	384	2	S25771	gas1 protein - mou
400	13	38.2	294	2	T23916	hypothetical prote	473	13	38.2	389	2	AD0722	glutamate dehydrog
401	13	38.2	294	2	T23682	hypothetical prote	474	13	38.2	393	2	A56947	dual specificity p
402	13	38.2	296	2	JT0548	killer toxin KHR p	475	13	38.2	394	2	A56115	dual specificity p
403	13	38.2	297	2	S06267	surface antigen H	476	13	38.2	395	2	AB0814	probable alcohol
404	13	38.2	300	2	T48147	hypothetical prote	477	13	38.2	396	2	B49475	Desert hedgehog pr
405	13	38.2	301	2	T05493	pathogenesis-relat	478	13	38.2	400	2	B48613	surface protein SU
406	13	38.2	303	2	T40593	cytoplasmic dynein	479	13	38.2	401	2	B87684	hypothetical prote
407	13	38.2	303	2	H71277	probable DNA adeni	480	13	38.2	403	2	A13256	hypothetical prote
408	13	38.2	303	2	T46405	hypothetical prote	481	13	38.2	404	2	S58830	probable membrane
409	13	38.2	303	2	B88990	protein C36C5.13 [482	13	38.2	410	2	AB3546	antibutyraldehyde
410	13	38.2	309	2	S37934	hypothetical prote	483	13	38.2	411	2	A82760	cell division prot
411	13	38.2	311	2	H64483	5-methyltetrahydro	484	13	38.2	411	2	T17653	hypothetical prote
412	13	38.2	312	2	B75405	endopeptidase-rele	485	13	38.2	412	2	C38104	LFY floral meriste
413	13	38.2	313	2	AC0603	asparaginase (EC 3	486	13	38.2	415	2	C84829	hypothetical prote
414	13	38.2	314	1	A57126	dual specificity p	487	13	38.2	415	2	T43352	nuclear receptor N
415	13	38.2	314	2	B57126	dual specificity p	488	13	38.2	416	2	S65110	chitinase (EC 3.2.
416	13	38.2	314	2	A44437	regenerating liver	489	13	38.2	420	2	B38104	LFY floral meriste
417	13	38.2	317	2	T00500	probable elicitor	490	13	38.2	422	2	S35197	hypothetical prote
418	13	38.2	317	2	D86292	protein F7H2.14 [i	491	13	38.2	424	2	A38104	LFY floral meriste
419	13	38.2	321	2	D64820	probable asparagin	492	13	38.2	425	2	B84631	probable serine ca
420	13	38.2	323	2	S19650	cysteine proteinas	493	13	38.2	425	2	F70608	hypothetical prote
421	13	38.2	323	2	A39877	autocrine motility	494	13	38.2	426	2	F85574	probable symptom p
422	13	38.2	324	1	D42951	alkanal monooxygen	495	13	38.2	426	2	F90723	probable transport
423	13	38.2	324	2	S47432	cathepsin L (EC 3.	496	13	38.2	430	2	JC4720	toxin co-regulated
424	13	38.2	324	2	C66582	hypothetical prote	497	13	38.2	433	1	S24353	proteaseome 26S sub
425	13	38.2	324	2	D72041	conserved hypotnet	498	13	38.2	433	2	S53709	MS1 protein homol
426	13	38.2	325	2	T27688	hypothetical prote	499	13	38.2	433	2	S39349	hypothetical prote
427	13	38.2	326	4	S61652	hypothetical prote	500	13	38.2	433	2	F86163	hypothetical prote
428	13	38.2	327	2	T41463	conserved hypotnet	501	13	38.2	434	2	S50865	avermectin-sensiti
429	13	38.2	329	2	S74227	cathepsin K (EC 3.	502	13	38.2	435	2	T20152	hypothetical prote
430	13	38.2	330	2	T25323	hypothetical prote	503	13	38.2	436	2	B81033	N-acetylglutamate
431	13	38.2	335	2	JE0115	zinc-finger protei	504	13	38.2	436	2	A81977	probable amino-ac
432	13	38.2	339	2	F47301	VirB11 homolog - B	505	13	38.2	437	2	S11497	branched-chain ami
433	13	38.2	341	1	T22930	phosphoprotein pho	506	13	38.2	437	2	H82295	branched chain ami
434	13	38.2	341	2	C70611	hemagglutinin prote	507	13	38.2	438	2	I38946	melanoma ubiquitou
435	13	38.2	344	2	JQ1643	hemagglutinin - in	508	13	38.2	440	2	D69596	branched-chain ami
436	13	38.2	346	1	M0H6D	N4-(beta-N-acetyl	509	13	38.2	440	2	F97098	branched-chain ami
437	13	38.2	347	2	S12955	calcium channel pr	510	13	38.2	445	2	S37779	porin precursor, m
438	13	38.2	348	2	T14141	NMDH2 dehydrogenas	511	13	38.2	445	2	C69596	branched-chain ami
439	13	38.2	349	2	B81855	adenine glycosylas	512	13	38.2	445	2	T18129	hypothetical prote
440	13	38.2	349	2	D81088	N/G-specific adeni	513	13	38.2	446	1	C70016	allantoinase homol
441	13	38.2	353	2	T27800	hypothetical prote	514	13	38.2	449	2	S50546	adenosylhomocyste
442	13	38.2	354	2	S70595	NMDH2 dehydrogen	515	13	38.2	453	2	T09745	myb-related protei
443	13	38.2	355	1	ADZM	fructose-bisphosph	516	13	38.2	454	2	D75446	cytochrome p450-11
444	13	38.2	355	2	T51414	CONSTANS-like 1 -	517	13	38.2	455	2	T48973	cytochrome p450-11
445	13	38.2	358	2	T47550	fructose bisphosph	518	13	38.2	455	2	AG0308	leucochormate synt
446	13	38.2	358	2	A84781	probable fructose	519	13	38.2	457	2	I48206	triacylglycerol 11
447	13	38.2	361	2	S59531	manganese peroxida	520	13	38.2	459	2	JC7931	cellulohydroxylase
448	13	38.2	362	2	T05824	probable pectin ac	521	13	38.2	461	2	S21223	triacylglycerol 11
449	13	38.2	363	2	JC5536	C-Fringe protein 1	522	13	38.2	461	2	S50864	avermectin-sensiti
450	13	38.2	364	1	T18972	phosphoprotein pho	523	13	38.2	464	2	C46157	hexokinase (EC 2.7
451	13	38.2	364	2	C45574	40.7 kDa polypepti	524	13	38.2	465	2	A46157	hexokinase (EC 2.7
452	13	38.2	364	2	H86384	40.1k. hypothetical	525	13	38.2	465	2	I49693	glucokinase (EC 2.
453	13	38.2	366	2	H89025	protein F13A2.6 [i	526	13	38.2	465	2	A31810	hexokinase (EC 2.7
454	13	38.2	367	1	S24411	dual specificity p	527	13	38.2	465	2	S41084	triacylglycerol 11
455	13	38.2	367	1	S29090	dual specificity p	528	13	38.2	466	2	B46157	hexokinase (EC 2.7
456	13	38.2	367	1	MOHU	alpha-2-HS-glycop	529	13	38.2	467	2	B89605	protein F18G5.2 [i
457	13	38.2	367	2	AH0936	glycerol dehydroge	530	13	38.2	471	2	S54460	hypothetical prote
458	13	38.2	367	2	S52265	dual specificity p	531	13	38.2	472	1	A26730	covalinhibitor precu
459	13	38.2	367	2	D82763	UDP-N-acetylglucos	532	13	38.2	474	2	B46746	glycine hydroxymet
460	13	38.2	367	2	JC6087	helix-loop-helix t	533	13	38.2	474	2	AF0759	cobyrinic acid A,C
461	13	38.2	368	2	H71600	reflin PFB1010w - m	534	13	38.2	475	2	A33366	glycine hydroxymet
462	13	38.2	371	2	T07938	probable acetyl-Co	535	13	38.2	476	2	A71302	conserved hypotnet
463	13	38.2	372	1	OPJGAP	lignin peroxidase	536	13	38.2	479	2	T49871	peroxisomal Ca-dep
464	13	38.2	372	2	B32322	lignin peroxidase	537	13	38.2	482	2	T48397	S-receptor kinase-
465	13	38.2	372	2	JH0156	lignin peroxidase	538	13	38.2	487	1	S71770	calcium-dependent
466	13	38.2	372	2	A53050	brefeldin A estera	539	13	38.2	487	2	A43758	atomatic-L-amino-a
467	13	38.2	377	2	T04213	heat shock transcr	540	13	38.2	488	2	D70876	probable polyketid

541	13	38.2	488	2	T30602	hypothetical prote	614	13	38.2	650	2	S44806	F10E9.6 protein -
542	13	38.2	489	2	H69059	dihydroxy-acid deh	615	13	38.2	654	2	C87791	protein B0207.12 (
543	13	38.2	494	2	C70940	probable cobo prot	616	13	38.2	656	2	D84359	thermosome subunit
544	13	38.2	495	2	T20754	hypothetical prote	617	13	38.2	655	2	E29826	hypothetical 72.4K
545	13	38.2	496	2	T33456	hypothetical prote	618	13	38.2	657	2	T52460	hypothetical prote
546	13	38.2	498	2	S12061	hexokinase (EC 2.7	619	13	38.2	659	2	A57532	Na+-dependent nucl
547	13	38.2	498	2	S27849	variant surface gl	620	13	38.2	663	2	T30621	hypothetical prote
548	13	38.2	504	2	T01603	hypothetical prote	621	13	38.2	664	2	S60062	hevin precursor -
549	13	38.2	504	2	T01603	hypothetical prote	622	13	38.2	669	2	A46511	envelope protein -
550	13	38.2	504	2	T16526	T22111.12 protein	623	13	38.2	675	2	G84653	hypothetical prote
551	13	38.2	506	2	S37583	hypothetical prote	624	13	38.2	677	2	C42125	trophozoite cystei
552	13	38.2	506	2	T29968	RING finger protei	625	13	38.2	680	2	G84730	Mutator-like trans
553	13	38.2	509	2	A95985	probable sugar kin	626	13	38.2	684	2	T11749	transferrin - Acta
554	13	38.2	511	2	S58339	celluloglucan I sylv	627	13	38.2	691	2	T31552	hypothetical prote
555	13	38.2	513	1	TVHURP	ret finger protein	628	13	38.2	700	2	S38426	chaperonin 60 prec
556	13	38.2	517	2	T19962	hypothetical prote	629	13	38.2	713	2	A35502	major surface-labe
557	13	38.2	520	2	A96922	beta-glucosidase h	630	13	38.2	730	2	A75486	hypothetical prote
558	13	38.2	530	2	JC7979	cellulohydrolase	631	13	38.2	733	1	A46373	probable serine/th
559	13	38.2	536	2	T24000	hypothetical prote	632	13	38.2	738	2	D86345	hypothetical prote
560	13	38.2	538	2	S52472	cell fusion protei	633	13	38.2	746	2	S52770	subtilisin-like pr
561	13	38.2	539	2	A29923	carboxylesterase (634	13	38.2	757	2	JC7519	subtilisin-like se
562	13	38.2	539	2	T46132	hypothetical prote	635	13	38.2	757	2	T16609	hypothetical prote
563	13	38.2	540	2	A31584	carboxylesterase (636	13	38.2	764	2	H83879	3-oxoacyl-(acyl)-ca
564	13	38.2	541	2	T33583	hypothetical prote	637	13	38.2	768	2	B41029	integrin beta-8 ch
565	13	38.2	541	2	B82342	probable transcrip	638	13	38.2	773	1	JH0609	protein-tyrosine-p
566	13	38.2	543	2	T27000	hypothetical prote	639	13	38.2	775	2	T00962	hypothetical prote
567	13	38.2	545	2	T49931	hypothetical prote	640	13	38.2	781	2	S43534	integrin beta3 - c
568	13	38.2	549	2	JX0054	carboxylesterase (641	13	38.2	795	2	T34468	hypothetical prote
569	13	38.2	554	2	A39060	carboxylesterase (642	13	38.2	796	1	JC1285	protein-tyrosine-p
570	13	38.2	557	2	A47162	cholesterase B (E	643	13	38.2	799	2	A38308	integrin beta-5 ch
571	13	38.2	557	2	A48434	variant-specific s	644	13	38.2	802	1	A36065	protein-tyrosine-p
572	13	38.2	560	1	F69059	arginine-tRNA liga	645	13	38.2	807	2	C85025	hypothetical prote
573	13	38.2	561	2	S71597	carboxylesterase (646	13	38.2	814	2	T49207	receptor kinase-1i
574	13	38.2	561	2	S62788	carboxylesterase (647	13	38.2	829	1	A47373	protein-tyrosine-p
575	13	38.2	561	2	C64459	dihydroxy-acid deh	648	13	38.2	832	2	JC8051	protein tyrosine p
576	13	38.2	562	2	A55281	carboxylesterase (649	13	38.2	840	2	S41218	HRI protein - yea
577	13	38.2	563	1	VCMMW7	env polypeptide -	650	13	38.2	843	1	JQ2229	DNA-directed DNA p
578	13	38.2	564	2	S08199	cytochrome-c3 hydr	651	13	38.2	848	1	T02053	S-receptor kinase
579	13	38.2	565	2	S10367	carboxylesterase (652	13	38.2	852	2	S28415	guanine nucleotide
580	13	38.2	566	2	S19307	carboxylesterase (653	13	38.2	852	2	T28790	hypothetical prote
581	13	38.2	566	2	S69889	hemagglutinin prec	654	13	38.2	885	2	JC7988	3',5'-cyclic nucle
582	13	38.2	567	1	A41010	carboxylesterase (655	13	38.2	891	2	T21942	hypothetical prote
583	13	38.2	567	2	T08405	hypothetical prote	656	13	38.2	892	2	A41697	nitrate assimilati
584	13	38.2	569	2	S50771	hexose transport p	657	13	38.2	895	2	T23191	hypothetical prote
585	13	38.2	571	1	S30253	GABA transport pro	658	13	38.2	907	2	T27317	hypothetical prote
586	13	38.2	573	2	T27671	hypothetical prote	659	13	38.2	917	1	S15885	hexokinase (EC 2.7
587	13	38.2	579	2	T30635	hypothetical prote	660	13	38.2	917	2	JC2025	hexokinase (EC 2.7
588	13	38.2	581	2	T24393	hypothetical prote	661	13	38.2	931	2	T25148	hypothetical prote
589	13	38.2	582	1	VCVDAR	env polypeptide -	662	13	38.2	934	2	G86548	polymorphic outer
590	13	38.2	591	2	I48141	acroganin - guine	663	13	38.2	934	2	A72075	polymorphic outer
591	13	38.2	594	2	H48613	env polypeptide pr	664	13	38.2	950	2	F86222	hypothetical prote
592	13	38.2	596	2	T30498	probable ribonucle	665	13	38.2	951	2	T00260	hypothetical prote
593	13	38.2	601	2	D89711	protein F40E10.4 (666	13	38.2	952	2	D81593	polymorphic membra
594	13	38.2	601	2	T22025	hypothetical prote	667	13	38.2	962	2	JC5571	subtilisin-like pr
595	13	38.2	602	1	S25316	hydroxymethylgluta	668	13	38.2	966	2	E84053	penicillin-binding
596	13	38.2	602	2	T37254	acetylcholinestera	669	13	38.2	969	1	A39490	subtilisin-like pr
597	13	38.2	605	2	S48940	hypothetical prote	670	13	38.2	975	2	JC5570	subtilisin-like pr
598	13	38.2	609	2	JC7819	metallopeptidase	671	13	38.2	983	2	T39902	translational elonga
599	13	38.2	616	2	T29234	hypothetical prote	672	13	38.2	993	2	S35633	DNA-binding protei
600	13	38.2	618	2	G82340	phosphogluconate d	673	13	38.2	994	2	H96510	probable disease r
601	13	38.2	625	2	T10661	serine/threonine-s	674	13	38.2	1011	2	T50344	poly(A)+ RNA trans
602	13	38.2	626	2	B97358	glucose-inhibited	675	13	38.2	1012	2	T52000	poly(A)+ RNA trans
603	13	38.2	632	2	S58152	hypothetical prote	676	13	38.2	1016	2	T00375	hypothetical prote
604	13	38.2	632	2	T21602	hypothetical prote	677	13	38.2	1046	2	A26838	prestalk protein p
605	13	38.2	633	1	T24898	matrix glycoprotei	678	13	38.2	1060	2	T30823	bumetanide sensiti
606	13	38.2	634	2	GERTX1	hypothetical prote	679	13	38.2	1073	2	E83051	cardanoylphosphate
607	13	38.2	635	2	S76371	hypothetical prote	680	13	38.2	1074	2	JC5928	semaphorin F precu
608	13	38.2	636	2	H87789	protein G34G6.1 (i	681	13	38.2	1082	2	T31112	Aspase 2 (EC 3.6.1
609	13	38.2	638	2	T26490	hypothetical prote	682	13	38.2	1087	2	S02035	period clock prote
610	13	38.2	640	1	HHKW7A	dnak-type molecula	683	13	38.2	1093	2	T50652	Ap-3 complex beta3
611	13	38.2	640	1	T07923	acetyl-CoA carboxy	684	13	38.2	1094	2	T50651	Ap3-complex beta-3
612	13	38.2	640	2	AB2251	glucose inhibited	685	13	38.2	1101	2	T16840	hypothetical prote
613	13	38.2	640	2	S62747	homeotic protein A	686	13	38.2	1105	2	C87622	TonB-dependent rec

687	13	38.2	1112	2	H97050	DNA Polymerase III	760	13	38.2	2143	2	JH0427	voltage-dependent
688	13	38.2	1129	2	T42732	A-kinase anchoring	761	13	38.2	2143	2	G96595	hypothetical prote
689	13	38.2	1149	2	I38006	M130 antigen precu	762	13	38.2	2166	2	S11339	calcium channel pr
690	13	38.2	1151	2	I38004	M130 antigen precu	763	13	38.2	2168	2	T30171	ninein - mouse
691	13	38.2	1156	2	I38005	M130 antigen precu	764	13	38.2	2171	2	S05054	calcium channel a1
692	13	38.2	1159	2	I38465	probable potassium	765	13	38.2	2180	2	T29764	hypothetical prote
693	13	38.2	1160	2	T13713	betalactam protein - fr	766	13	38.2	2188	2	A70984	probable polyketid
694	13	38.2	1162	2	T21557	hypothetical prote	767	13	38.2	2195	2	T34264	hypothetical prote
695	13	38.2	1166	2	T13958	synGAP-bl protein	768	13	38.2	2201	2	S73014	polyketide synthas
696	13	38.2	1170	2	I45914	integrin alpha 2 s	769	13	38.2	2204	1	RKNZNV	genome polypotein
697	13	38.2	1176	2	T49482	hypothetical prote	770	13	38.2	2220	2	A45290	calcium channel pr
698	13	38.2	1180	2	A35854	integrin alpha-1 c	771	13	38.2	2233	2	T28669	surface protein 51
699	13	38.2	1181	2	A33998	integrin alpha-2 c	772	13	38.2	2240	2	T37057	probable multi-dom
700	13	38.2	1205	2	T41987	hypothetical prote	773	13	38.2	2257	2	D86483	protein F5J5.19 [i
701	13	38.2	1242	2	S14201	probable adenylate	774	13	38.2	2257	2	T09538	acetyl-CoA carboxy
702	13	38.2	1249	2	T14270	Ras-GTPase activat	775	13	38.2	2261	2	T07084	probable acetyl-Co
703	13	38.2	1265	2	S57968	Ran-binding protei	776	13	38.2	2304	2	T07920	polyketide synthas
704	13	38.2	1271	2	T24008	transferrin-like p	777	13	38.2	2359	2	E86483	notch protein - Af
705	13	38.2	1274	2	T10729	cysteine rich prot	778	13	38.2	2518	2	A12140	notch protein homo
706	13	38.2	1274	2	T42017	ras GTPase-activat	779	13	38.2	2524	2	A35844	notch-1 protein -
707	13	38.2	1293	2	T14259	furin (EC 3.4.21.7	780	13	38.2	2531	2	A46019	polyketide synthas
708	13	38.2	1299	2	T43251	connectin 3B - chl	781	13	38.2	2531	2	A46019	notch protein homo
709	13	38.2	1323	2	PM0568	protein let-23 [im	782	13	38.2	2543	2	F69679	hypothetical prote
710	13	38.2	1323	2	E88257	RNA2 polypeptide -	783	13	38.2	2555	2	A40043	variant-specific s
711	13	38.2	1324	2	S06187	calcium channel a1	784	13	38.2	2610	2	T20968	hypothetical prote
712	13	38.2	1331	2	S05011	protein F21D18.16	785	13	38.2	2706	2	T28155	fibroblast-in-2 precu
713	13	38.2	1340	2	D96521	xanthine dehydroge	786	13	38.2	2844	2	S28291	hypothetical prote
714	13	38.2	1359	2	T10235	alpha-glycan synth	787	13	38.2	2918	2	A54105	hypothetical prote
715	13	38.2	1369	2	T13433	protein-tyrosine k	788	13	38.2	2946	2	T15840	protein HMW1 - ye
716	13	38.2	1374	2	S70712	hypothetical prote	789	13	38.2	3161	2	T30342	lysine-specific bio
717	13	38.2	1391	2	D70634	probable polyketid	790	13	38.2	3163	2	AB0223	probable polyketid
718	13	38.2	1402	2	D70634	probable membrane	791	13	38.2	3163	2	T17440	Ran-binding protei
719	13	38.2	1407	2	S59823	protein ZC84.1 [im	792	13	38.2	3224	1	S58884	reelin precursor -
720	13	38.2	1416	2	E88550	polyketide synthas	793	13	38.2	3461	2	S58870	hypothetical prote
721	13	38.2	1446	2	S73013	hypothetical prote	794	13	38.2	3507	2	T34513	trithorax protein
722	13	38.2	1461	2	T26327	breast cancer tumo	795	13	38.2	3828	2	T13857	mycosubtilin synth
723	13	38.2	1484	2	T42632	polyketide synthas	796	13	38.2	3971	2	T44806	connectin/titin -
724	13	38.2	1488	2	AG2136	probable ppas prot	797	13	38.2	4162	2	T42215	zonadhesin - mouse
725	13	38.2	1488	2	C70984	polyketide synthas	798	13	38.2	5376	2	T42215	polyketide synthas
726	13	38.2	1489	2	S73015	probable ppas prot	799	13	38.2	6420	2	T30283	CAP5 protein - ant
727	13	38.2	1538	2	E70874	polyketide synthas	800	13	35.3	27	2	S55030	neurotoxin Bmk 41-
728	13	38.2	1540	2	H87203	polyketide synthas	801	12	35.3	35	2	A59356	metallothionein Mt
729	13	38.2	1565	2	AD2135	antigen containing	802	12	35.3	40	1	SMF	Kit-ligand (alterm
730	13	38.2	1567	2	T03730	hypothetical prote	803	12	35.3	40	2	B61194	hypothetical prote
731	13	38.2	1570	2	AC2012	hypothetical prote	804	12	35.3	41	2	B44010	proteinase inhibit
732	13	38.2	1587	2	AB2012	gene small optic 1	805	12	35.3	43	2	D84641	leech-derived tryp
733	13	38.2	1597	1	BYF8SL	protein F11C7.4 [i	806	12	35.3	45	1	XKPO2A	hypothetical prote
734	13	38.2	1597	2	T08428	hypothetical synthas	807	12	35.3	46	2	S50015	conotoxin-like pro
735	13	38.2	1722	2	E89753	hypothetical prote	808	12	35.3	52	2	T10299	ovomucoid, third d
736	13	38.2	1743	2	T26859	hypothetical synthas	809	12	35.3	54	2	A31436	ovomucoid, third d
737	13	38.2	1744	2	T17421	polyketide synthas	810	12	35.3	54	2	F61589	ovomucoid (PSTI-ty
738	13	38.2	1795	2	D97312	polyketide synthas	811	12	35.3	54	2	F61589	ovomucoid (PSTI-ty
739	13	38.2	1815	2	S73021	polyketide synthas	812	12	35.3	54	2	F61589	ovomucoid (PSTI-ty
740	13	38.2	1822	2	F87203	probable polyketid	813	12	35.3	54	2	F61588	ovomucoid, third d
741	13	38.2	1827	2	B70984	genome polypotein	814	12	35.3	54	2	F31439	ovomucoid, third d
742	13	38.2	1844	1	RRPPTM	hypothetical prote	815	12	35.3	54	2	T31446	ovomucoid, third d
743	13	38.2	1844	2	S01956	polyketide synthas	816	12	35.3	54	2	C31447	ovomucoid (PSTI-ty
744	13	38.2	1871	2	A87204	probable ppas prot	817	12	35.3	54	2	D61589	ovomucoid (PSTI-ty
745	13	38.2	1876	2	C70749	agrin - rat	818	12	35.3	54	2	B61492	ovomucoid (PSTI-ty
746	13	38.2	1959	1	AGRT	notch4 - mouse	819	12	35.3	54	2	G31442	ovomucoid, third d
747	13	38.2	1964	2	T09059	acetyl-CoA carboxy	820	12	35.3	54	2	A61492	ovomucoid (PSTI-ty
748	13	38.2	1978	2	T07081	probable vitellogen	821	12	35.3	54	2	H31441	ovomucoid, third d
749	13	38.2	1984	2	T13171	hypothetical prote	822	12	35.3	54	2	H31442	ovomucoid, third d
750	13	38.2	1985	2	S19151	hypothetical prote	823	12	35.3	54	2	T31444	ovomucoid, third d
751	13	38.2	2004	2	T30185	hypothetical prote	824	12	35.3	54	2	B31444	ovomucoid, third d
752	13	38.2	2014	2	T21560	hypothetical prote	825	12	35.3	54	2	F31436	ovomucoid, third d
753	13	38.2	2027	2	S60123	hypothetical prote	826	12	35.3	54	2	H31436	ovomucoid (PSTI-ty
754	13	38.2	2039	2	S64540	probable calcium c	827	12	35.3	54	2	H31445	ovomucoid, third d
755	13	38.2	2056	2	G88564	protein R10E11.1 [828	12	35.3	54	2	F61494	ovomucoid (PSTI-ty
756	13	38.2	2108	2	H70819	probable polyketid	829	12	35.3	54	2	F31447	ovomucoid, third d
757	13	38.2	2129	2	T28658	polyketide synthas	830	12	35.3	54	2	F61492	ovomucoid (PSTI-ty
758	13	38.2	2126	2	E70522	probable polyketid	831	12	35.3	54	2	H61492	ovomucoid (PSTI-ty
759	13	38.2	2139	2	A44467	voltage-dependent	832	12	35.3	54	2	E31447	ovomucoid, third d

833	12	35.3	54	2	B31443	ovomucoid, third d	906	12	35.3	111	2	B85888	detox protein (imp
834	12	35.3	54	2	A31439	ovomucoid, third d	907	12	35.3	112	2	JC7528	guanylate cyclase
835	12	35.3	54	2	G61492	ovomucoid (PSTI-ty	908	12	35.3	111	1	XLHU	collipase precursor
836	12	35.3	54	2	B31436	ovomucoid, third d	909	12	35.3	112	2	A46717	collipase precursor
837	12	35.3	54	2	B31436	ovomucoid, third d	910	12	35.3	112	2	A64448	hypothetical prote
838	12	35.3	55	2	B31442	ovomucoid, third d	911	12	35.3	113	2	F95248	conserved hypotnet
839	12	35.3	56	2	D31445	ovomucoid, third d	912	12	35.3	113	2	C98113	conserved hypotnet
840	12	35.3	56	2	A31446	ovomucoid, third d	913	12	35.3	115	2	S66924	probable membrane
841	12	35.3	56	2	C61588	ovomucoid (PSTI-ty	914	12	35.3	115	2	G95059	hypothetical prote
842	12	35.3	56	2	G98233	hypothetical prote	915	12	35.3	115	2	B96664	probable RING zinc
843	12	35.3	56	2	A25356	bdellin B-3 - medi	916	12	35.3	115	2	S45920	hypothetical prote
844	12	35.3	59	2	A47218	carboxyl ester lip	917	12	35.3	119	1	FSNU3M	phospholipase A2 (
845	12	35.3	60	2	I61886	zinc finger protei	918	12	35.3	121	1	QOEC13	hypothetical 13..5K
846	12	35.3	60	2	B86123	hypothetical prote	919	12	35.3	121	2	PH1661	Ig heavy chain V r
847	12	35.3	62	2	S62862	toxin III-8 - Tily	920	12	35.3	121	2	T02926	acyl carrier prote
848	12	35.3	62	2	S62866	toxin III-8 - Tily	921	12	35.3	121	2	A98089	hypothetical prote
849	12	35.3	62	2	B39510	toxin III-8 - Braz	922	12	35.3	121	2	C85934	hypothetical prote
850	12	35.3	62	2	S54336	metallothionein-2a	923	12	35.3	121	2	QO1058	DNA-binding protei
851	12	35.3	62	2	S54335	metallothionein-2c	924	12	35.3	122	2	T18564	hypothetical prote
852	12	35.3	64	2	S66481	hypothetical prote	925	12	35.3	123	2	S30530	Ig heavy chain V r
853	12	35.3	66	2	E70083	hypothetical prote	926	12	35.3	124	2	T01251	hypothetical prote
854	12	35.3	67	2	JN0378	neurotoxin Os-3 -	927	12	35.3	125	1	NFBO1	oxytocin / neuroph
855	12	35.3	67	2	S73050	probable acetyltra	928	12	35.3	125	1	NFHU1	oxytocin / neuroph
856	12	35.3	72	1	N2EPIJ	long neurotoxin 1	929	12	35.3	125	1	NFPG1	oxytocin / neuroph
857	12	35.3	72	1	F82825	hypothetical prote	930	12	35.3	125	1	NFR1	oxytocin / neuroph
858	12	35.3	73	1	N2OH2	long neurotoxin 2	931	12	35.3	125	1	NFSH1	oxytocin / neuroph
859	12	35.3	77	2	B82617	hypothetical prote	932	12	35.3	125	2	A43755	oxytocin / neuroph
860	12	35.3	78	2	E70891	hypothetical prote	933	12	35.3	128	2	S26786	Ig heavy chain V r
861	12	35.3	82	2	S70807	hypothetical prote	934	12	35.3	128	2	QO2180	hypothetical 14..8K
862	12	35.3	84	1	JU0152	acrosin/trypsin in	935	12	35.3	128	2	JU01739	hypothetical 14..7K
863	12	35.3	85	2	D34770	ORP4 protein - sai	936	12	35.3	128	2	S27917	hypothetical prote
864	12	35.3	86	2	I61885	zinc finger protei	937	12	35.3	129	2	S01661	lysozyme (EC 3.2.1
865	12	35.3	89	2	D43692	TfB protein - rabb	938	12	35.3	129	2	S07435	lysozyme (EC 3.2.1
866	12	35.3	89	2	S71555	proteinase inhibit	939	12	35.3	130	2	G71212	hypothetical prote
867	12	35.3	89	2	VQ2361	wheat aluminum ind	940	12	35.3	131	2	I40656	hypothetical prote
868	12	35.3	92	2	T10783	protein mcl - Mont	941	12	35.3	132	2	F64487	hypothetical prote
869	12	35.3	93	2	T06470	probable chitinase	942	12	35.3	133	2	T15955	hypothetical prote
870	12	35.3	94	2	T03285	anther-specific pr	943	12	35.3	133	2	T37311	hypothetical prote
871	12	35.3	96	1	XLHOB	collipase B precurs	944	12	35.3	133	2	JH0270	hypothetical prote
872	12	35.3	97	1	KRGTHM	keratin high-sulfu	945	12	35.3	133	2	JE0311	chondromodulin II
873	12	35.3	97	1	KRSHH2	keratin high-sulfu	946	12	35.3	135	2	AF0865	conserved hypotnet
874	12	35.3	97	2	T42335	hypothetical prote	947	12	35.3	135	2	S36197	hypothetical prote
875	12	35.3	97	2	JU0041	hypothetical prote	948	12	35.3	135	2	G96691	hypothetical prote
876	12	35.3	97	2	D82789	hypothetical prote	949	12	35.3	136	2	T36241	hypothetical prote
877	12	35.3	97	2	C95984	hypothetical prote	950	12	35.3	137	1	JC4233	lysozyme (EC 3.2.1
878	12	35.3	98	1	S25283	amebapore A precu	951	12	35.3	137	2	T37213	hypothetical prote
879	12	35.3	98	1	KRSHH3	keratin high-sulfu	952	12	35.3	137	2	G84174	hypothetical prote
880	12	35.3	98	1	KRSHH4	keratin high-sulfu	953	12	35.3	138	2	AD0281	conserved hypotnet
881	12	35.3	98	2	S26913	Ig heavy chain V r	954	12	35.3	138	2	S34725	regulation protein
882	12	35.3	98	2	I47083	BLIIB2 high-sulfu	955	12	35.3	139	2	T29955	hypothetical prote
883	12	35.3	98	2	I47086	BLIIB2 high-sulfu	956	12	35.3	140	2	JC5003	lysozyme (EC 3.2.1
884	12	35.3	98	2	B86719	hypothetical prote	957	12	35.3	140	2	T49347	hypothetical prote
885	12	35.3	100	2	F95897	hypothetical prote	958	12	35.3	140	2	T05525	hypothetical prote
886	12	35.3	101	2	F75500	hypothetical prote	959	12	35.3	141	2	T33710	hypothetical prote
887	12	35.3	101	2	T26641	hypothetical prote	960	12	35.3	141	2	PC1294	trophozoite surfac
888	12	35.3	101	2	F72544	hypothetical prote	961	12	35.3	142	2	AC1934	hypothetical prote
889	12	35.3	102	2	PH1249	Ig heavy chain V r	962	12	35.3	144	2	A89472	protein B0294..1 [i
890	12	35.3	102	2	B89864	conserved hypotnet	963	12	35.3	145	2	S74282	hypothetical prote
891	12	35.3	102	2	S50530	hypothetical prote	964	12	35.3	145	2	S02027	hemoglobin beta ch
892	12	35.3	102	2	A72604	hypothetical prote	965	12	35.3	145	2	G96581	F511..18 [imported
893	12	35.3	104	2	I50808	neuropeptide Y pre	966	12	35.3	145	2	T34539	hypothetical prote
894	12	35.3	104	2	T43968	hypothetical prote	967	12	35.3	146	2	B82568	conserved hypotnet
895	12	35.3	104	2	F45681	orf 61.2 - phage T	968	12	35.3	146	2	B95363	protein [imported
896	12	35.3	106	2	F72043	conserved hypotnet	969	12	35.3	147	2	S34148	thyrocytopin beta c
897	12	35.3	106	2	C66580	Ct57 hypotnetical	970	12	35.3	147	2	T47192	Ig heavy chain var
898	12	35.3	107	2	T51038	hypotnetical prote	971	12	35.3	148	2	H71644	hypotnetical prote
899	12	35.3	107	2	A82642	conserved plasmid	972	12	35.3	148	2	B49727	DNA-binding protei
900	12	35.3	108	2	E70931	hypotnetical prote	973	12	35.3	148	2	T21334	hypotnetical prote
901	12	35.3	108	2	AD0715	conserved hypotnet	974	12	35.3	149	2	JC5396	helix-loop-helix p
902	12	35.3	108	2	JN0368	asparaginopepsin I	975	12	35.3	151	2	T36314	hypotnetical prote
903	12	35.3	109	2	S54446	prepilin-like prot	976	12	35.3	151	1	KRSHHC	keratin high-sulfu
904	12	35.3	111	2	G91043	detox protein (imp	977	12	35.3	152	2	S57427	cysteine proteinas
905	12	35.3	111	2	H65020	ccha protein - Bsc	978	12	35.3	152	2	I47111	high-sulfur wool m

979 12 35.3 152 2 147109 high-sulfur wool m
980 12 35.3 152 2 147112 finger protein wool m
981 12 35.3 152 2 C39364 hypothetical prote
982 12 35.3 153 2 C70958 hypothetical prote
983 12 35.3 153 2 A96751 B6 protein - human
984 12 35.3 154 2 S36584 DNA-binding protei
985 12 35.3 154 2 A49727 gene Id1 protein -
986 12 35.3 155 1 S47524 trypsin/factor XII
987 12 35.3 155 1 T12M1 helix-loop-helix p
988 12 35.3 155 2 UC5395 hypothetical prote
989 12 35.3 155 2 D72761 trypsin-plasmin in
990 12 35.3 155 2 UC7732 trypsin-plasmin in
991 12 35.3 157 2 S09805 vasotocin / neurop
992 12 35.3 158 2 B34132 hypothetical prote
993 12 35.3 158 2 G86964 hypothetical prote
994 12 35.3 158 2 T26444 coagulation factor
995 12 35.3 159 2 I84615 hypothetical 17.8K
996 12 35.3 159 2 UC0136 Salp2R protein pre
997 12 35.3 159 2 JQ1777 early E3 18.5K gly
998 12 35.3 160 1 T33696 cysteine-rich exte
999 12 35.3 160 1 ERADA5
1000 12 35.3 160 2 D48232

ALIGNMENTS

RESULT 1

GSFP3 salivary glue protein sgs-3 - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster

C/Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 09-Jul-2004

C/Accession: A03329

R.Garfinkel, M.D.; Pruitt, R.E.; Meyerowitz, E.M.

J. Mol. Biol. 168, 765-789, 1983

A/Title: DNA sequences, gene regulation and modular protein evolution in the Drosophila

A/Reference number: A92904; MUID:83294545; PMID:6411930

A/Accession: A03329

A/Molecule type: DNA

A/Residues: 1-307 <GAR>

A/Cross-references: UNIPROT:P02840; UNIPARC:UPI00001245F4; GB:X01918; NID:98581; PIDN:CA

C/Comment: This protein is produced by third-instar larvae.

C/Genetics:

A/Gene: sgs-3

A/Cross-references: FlyBase:FBgn000373

A/Map position: 3L (68C)

A/Intons: 10/1

C/Superfamily: salivary glue protein

C/Keywords: salivary gland; tandem repeat

Query Match 52.9%; Score 18; DB 1; Length 307;
Best Local Similarity 16.7%; Pred. No. 1.3e-10;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15

Db 44 CTTTTTTTTTTC 55

RESULT 2

T24565 hypothetical protein T06C12.14 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T24565

R.Kelly, P.

Submitted to the EMBL Data Library, October 1996

A/Reference number: Z19908

A/Accession: T24565

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-245 <WIL>

A/Cross-references: UNIPROT:O18039; UNIPARC:UPI000007C269; EMBL:Z81116; PIDN:CAB03300.1;

A/Experimental source: clone T06C12
C/Genetics:
A/Gene: CESP:T06C12.14
A/Map position: 5
A/Intons: 33/3; 121/1
C/Superfamily: Caenorhabditis elegans hypothetical protein C49G7.3

Query Match 50.0%; Score 17; DB 2; Length 245;
Best Local Similarity 16.7%; Pred. No. 5.8e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15

Db 193 CASATTTTSSSTC 204

RESULT 3

D72630 hypothetical protein APE1499 - Aeropyrum pernix (strain K1)

C/Species: Aeropyrum pernix

C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C/Accession: D72630

R.Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, D.; Ki

DNA Res. 6, 83-101, 1999

A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A/Reference number: A72450; MUID:99310339; PMID:10382966

A/Accession: D72630

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-152 <KAW>

A/Cross-references: UNIPROT:Q9YBV0; UNIPARC:UPI000005DF8F; DBJ:AF000061; NID:G5104821; I

A/Experimental source: strain K1

C/Genetics:

A/Gene: APE1499

Query Match 47.1%; Score 16; DB 2; Length 152;
Best Local Similarity 16.7%; Pred. No. 2.5e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15

Db 137 CSSTATTTLSTRC 148

RESULT 4

T31838 hypothetical protein T05B4.12 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C/Accession: T31838

R.Bradsaw, H.

Submitted to the EMBL Data Library, July 1997

A/Description: The sequence of C. elegans cosmid T05B4.

A/Reference number: Z21092

A/Accession: T31838

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-244 <BRA>

A/Cross-references: UNIPROT:O16421; UNIPARC:UPI000007AC7E; EMBL:AF016445; PIDN:AA69052.1

A/Experimental source: strain Bristol N2; clone T05B4

C/Genetics:

A/Gene: CESP:T05B4.12

A/Map position: 5

A/Intons: 30/3; 116/1

C/Superfamily: Caenorhabditis elegans hypothetical protein C49G7.3

Query Match 47.1%; Score 16; DB 2; Length 244;
Best Local Similarity 16.7%; Pred. No. 2.8e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15

Db 203 CTSYADSTSC 214

RESULT 5

T31841
hypothetical protein T05B4.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C/Accession: T31841

R/Bradshaw, H.
submitted to the EMBL Data Library, July 1997

A/Description: The sequence of C. elegans cosmid T05B4.

A/Reference number: Z21092

A/Accession: T31841

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-248 <BRA>

A/Cross-references: UNIPROT:O16424; UNIPARC:UPI000007P78C; EMBL:AF016445; PIDN:AA69054.

A/Experimental source: strain Bristol N2; clone T05B4

C/Genetics:

A/Gene: CESP:T05B4.3

A/Map position: 5

A/Intons: 30/3; 118/1

C/Superfamily: Caenorhabditis elegans hypothetical protein C49G7.3

Query Match

Best Local Similarity 47.1%; Score 16; DB 2; Length 248;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

Db 207 CTSFADSSSC 218

RESULT 6

T21029
hypothetical protein F16H6.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T21029

R/Mathews, L.
submitted to the EMBL Data Library, November 1996

A/Reference number: Z19362

A/Accession: T21029

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-257 <WIL>

A/Cross-references: UNIPROT:Q9XV90; UNIPARC:UPI0000075F8F; EMBL:Z81506; PIDN:CAB04129.1;

A/Experimental source: clone F16H6

C/Genetics:

A/Gene: CESP:F16H6.3

A/Map position: 5

A/Intons: 39/3; 123/1

C/Superfamily: Caenorhabditis elegans hypothetical protein C49G7.3

Query Match

Best Local Similarity 47.1%; Score 16; DB 2; Length 257;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

Db 195 CTNATTTQASC 206

RESULT 7

T52565
glutathione synthase (BC 6.3.2.3) [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 09-Jul-2004

C/Accession: T52565

R/Skipsey, M.; Andrews, C.J.; Townson, J.K.; Jepson, I.; Edwards, R.

submitted to the EMBL Data Library, July 1999

A/Description: Isolation of cDNA and genomic clones of glutathione synthetase containing

A/Reference number: Z26117

A/Accession: T52565

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-539 <SKI>

A/Cross-references: UNIPROT:P46416; UNIPARC:UPI000016DB47; EMBL:AJ243813; PIDN:CAB51027.

A/Experimental source: cultivar Columbia; seedling; whole plant

C/Genetics:

A/Gene: gsh2

C/Superfamily: glutathione synthase

C/Keywords: ligase

Query Match

Best Local Similarity 47.1%; Score 16; DB 2; Length 539;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

Db 5 CSSLVSSSSSTC 16

RESULT 8

A48579
trophozoite surface protein TSP11 - Giardia lamblia

C/Species: Giardia lamblia

C/Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C/Accession: A48579

R/B, P.L.; Khanna, K.K.; Manning, P.A.; Mayrhofer, G.

Mol. Biochem. Parasitol. 58, 247-257, 1993

A/Title: A gene encoding a 69-kilodalton major surface protein of Giardia intestinalis t:

A/Reference number: A48579; MUID:93241215; PMID:8479449

A/Contents: Ad-1

A/Accession: A48579

A/Status: preliminary

A/Molecule type: nucleic acid

A/Residues: 1-667 <EY1>

A/Cross-references: UNIPROT:Q03185; UNIPARC:UPI0000137726; GB:M95814; NID:G159106; PID:G:

A/Note: sequence extracted from NCBI backbone (NCBI:130056, NCBI:130058)

Query Match

Best Local Similarity 47.1%; Score 16; DB 2; Length 667;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

Db 449 CTSTTARTVATC 460

RESULT 9

T51024
related to C2H2 zinc finger transcription factor D-Spl [imported] - Neurospora crassa

N/Alternate names: Protein B7F21.50

C/Species: Neurospora crassa

C/Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004

C/Accession: T51024

R/Schulte, U.; Aign, V.; Hohnsels, J.; Brandt, P.; Fattmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, July 2000

A/Reference number: Z25286

A/Accession: T51024

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-770 <SCH>

A/Cross-references: UNIPROT:Q9P319; UNIPARC:UPI000017B4E0; EMBL:AL389901; GSPDB:GN00116;

A/Experimental source: BAC clone B7F21; strain OR7A

C/Genetics:

A/Gene: NCSP:B7F21.50

A/Map position: 6

A/Intons: 117/1

Query Match

Best Local Similarity 47.1%; Score 16; DB 2; Length 770;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
 Db 92 CSATTAAAGTAC 103

RESULT 10

T31097
 chitin synthase (EC 2.4.1.16) Chs1 - fungus (*Filobasidium floriforme*)
 C/Species: *Filobasidiella neoformans*, *Cryptococcus neoformans*
 C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
 C/Accession: T31097
 R/Specht, C.A.
 submitted to the EMBL Data Library, August 1997
 A/Description: Chs1, a class IV chitin synthase of *Cryptococcus neoformans*.
 A/Reference number: Z20980
 A/Accession: T31097
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1041 <SPE>
 A/Cross-references: UNIPROT:O13356; UNIPARC:UPI00001278CD; EMBL:AF021318; NID:G2444456;
 A/Experimental source: strain H99
 C/Genetics:
 A/Gene: Chs1
 A/Introns: 152/1; 556/1; 634/3; 922/2
 C/Function:
 A/Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucose
 A/Pathway: chitin biosynthesis
 A/Note: class IV chitin synthase
 C/Superfamily: chitin synthase chs4
 C/Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 47.1%; Score 16; DB 2; Length 1041;
 Best Local Similarity 16.7%; Pred. No. 3.7e-07;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
 Db 212 CNSTSSSSGTSC 223

RESULT 11

AC0416
 probable prepilin peptidase dependent protein D precursor ppdd [imported] - *Yersinia pestis*
 C/Species: *Yersinia pestis*
 C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
 C/Accession: AC0416
 R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M.B.;
 demo-Terraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
 Nature 413, 523-527, 2001
 A/Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A/Reference number: AB0001; MUID:21470413; PMID:11586360
 A/Accession: AC0416
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-154 <KUR>
 A/Cross-references: UNIPROT:Q8ZB15; UNIPARC:UPI00000DC98A; GB:AL590842; PIDN:CAC92655.1;
 C/Genetics:
 A/Gene: ppdd
 C/Superfamily: gonococcal fimbrial protein

Query Match 44.1%; Score 15; DB 2; Length 154;
 Best Local Similarity 16.7%; Pred. No. 1.2e-05;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
 Db 129 CTATDSAMTSC 140

RESULT 12

B41378
 cytochrome c553i precursor - *Paracoccus denitrificans*

C/Species: *Paracoccus denitrificans*
 C/Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 09-Jul-2004
 C/Accession: B41378
 R/Rae, J.; Reijnders, W.N.M.; Van Spanning, R.J.M.; Harms, N.; Oltmann, L.F.; Stouthamer,
 J. Bacteriol. 173, 6971-6979, 1991
 A/Title: Isolation, sequencing, and mutagenesis of the gene encoding cytochrome c-553i of
 A/Reference number: A41378; MUID:92041583; PMID:1657873
 A/Accession: B41378
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-226 <RAS>
 A/Cross-references: UNIPROT:P29967; UNIPARC:UPI0000126C5E; GB:M75583; NID:G150576; PIDN:
 C/Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-226/Product: cytochrome c553i #status predicted <MAT>
 F:115-193/Domain: cytochrome c6 homology <CYC>
 F:125,128/Binding site: heme (Cys) (covalent) #status predicted
 F:129,173/Binding site: heme iron (His, Met) (axial ligands) #status predicted

Query Match 44.1%; Score 15; DB 2; Length 226;
 Best Local Similarity 16.7%; Pred. No. 1.3e-05;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
 Db 16 CAASAIAGTALC 27

RESULT 13

T21786
 hypothetical protein F35E8.9 - *Caenorhabditis elegans*
 C/Species: *Caenorhabditis elegans*
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C/Accession: T21786
 R/Gardner, A.
 submitted to the EMBL Data Library, November 1996
 A/Reference number: Z19472
 A/Accession: T21786
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-248 <WIL>
 A/Cross-references: UNIPROT:O45450; UNIPARC:UPI0000077E4A; EMBL:Z81529; PIDN:CAB04292.1;
 A/Experimental source: clone F35E8
 C/Genetics:
 A/Gene: CESP:F35E8.9
 A/Map position: 5
 A/Introns: 33/3; 121/1
 C/Superfamily: *Caenorhabditis elegans* hypothetical protein C49G7.3

Query Match 44.1%; Score 15; DB 2; Length 248;
 Best Local Similarity 16.7%; Pred. No. 1.3e-05;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
 Db 207 CSSYIADSTAC 218

RESULT 14

A84746
 probable NAM (no apical meristem)-like protein [imported] - *Arabidopsis thaliana*
 C/Species: *Arabidopsis thaliana* (mouse-ear cress)
 C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C/Accession: A84746
 R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 Nature 402, 761-768, 1999
 A/Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A/Reference number: A84420; MUID:20083487; PMID:10617197
 A/Accession: A84746
 A/Status: preliminary
 A/Molecule type: DNA

A:Residues: 1-268 <STD>
A:Cross-references: UNIPROT:O22798; UNIPARC:UPI00000A7D54; GB:AE002093; NID:g2459430; PDB:1A2374.1
C:Species: Pisum sativum (garden pea)
A:Gene: At2g33480
A:Map position: 2

Query Match 44.1%; Score 15; DB 2; Length 268;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 230 CCSSSTSSSSVC 241

RESULT 15

S69193
probable finger protein Pszf1 - garden pea
C:Species: Pisum sativum (garden pea)
C>Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
R:Michael, A.J.; Hofer, J.M.I.; Ellis, T.H.N.
Plant Mol. Biol. 30, 1051-1058, 1996
A:Title: Isolation by PCR of a cDNA clone from pea petals with similarity to petunia and
A:Reference number: S69193; MUID:96270382; PMID:8639742
A:Accession: S69193
A:Molecule type: mRNA
A:Residues: 1-273 <MTC>
A:Cross-references: UNIPROT:Q41070; UNIPARC:UPI000009E1A7; EMBL:X87374; NID:g854399; PDB:1A2374.1
C:Superfamily: Arabidopsis thaliana hypothetical protein F12E4.290
C:Keywords: DNA binding; zinc finger
F:115-135/Region: zinc finger CCH motif
F:202-222/Region: zinc finger CCH motif

Query Match 44.1%; Score 15; DB 2; Length 273;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 29 CTAVNTVTSSC 40

RESULT 16

AH2744
conserved hypothetical protein Atcu1368 [imported] - Agrobacterium tumefaciens (strain C58)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
R:Wood, D.W.; Seubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, L.; Karp, P.; Kometani, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, S.; Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.; W.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AH2744
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-283 <KDR>
A:Cross-references: UNIPROT:Q8UPM9; UNIPARC:UPI00001645DD; GB:AE008688; PDB:1A2374.1
C:Genetics:
A:Gene: Atcu1368
A:Map position: circular chromosome

Query Match 44.1%; Score 15; DB 2; Length 283;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

DB 242 CSATSSDMAAC 253

RESULT 17

S57770
xyloglucan endo-1,4-beta-D-glucanase (EC 3.2.1.-) precursor (clone tXET-B2) - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 06-Dec-1996 #sequence_revision 07-Feb-1997 #text_change 09-Jul-2004
R:Arrowsmith, D.A.; de Silva, J.
Plant Mol. Biol. 28, 391-403, 1995
A:Title: Characterisation of two tomato fruit-expressed cDNAs encoding xyloglucan endo-1,4-beta-D-glucanase
A:Reference number: S57769; MUID:95359399; PMID:7632911
A:Accession: S57770
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-287 <ARR>
A:Cross-references: UNIPROT:Q43528; UNIPARC:UPI00000AAE06; EMBL:X82684; NID:g577067; PDB:1A2374.1
A:Note: The sequence from Fig. 2 is inconsistent with that from Fig. 5 in having 246-Glu
C:Superfamily: endoxyloglucan transferase
C:Keywords: glycosidase; hydrolase
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-287/Product: xyloglucan endo-transglycosylase #status predicted <MAT>

Query Match 44.1%; Score 15; DB 2; Length 287;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 217 CWSAATSTSSC 228

RESULT 18

S49812
xyloglucan endo-1,4-beta-D-glucanase (EC 3.2.1.-) precursor (clone tXET-B1) - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
R:Arrowsmith, D.A.; de Silva, J.
Plant Mol. Biol. 28, 391-403, 1995
A:Title: Characterisation of two tomato fruit-expressed cDNAs encoding xyloglucan endo-1,4-beta-D-glucanase
A:Reference number: S57769; MUID:95359399; PMID:7632911
A:Accession: S57769
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-289 <ARR>
A:Cross-references: UNIPROT:Q43527; UNIPARC:UPI00000A4E68; EMBL:X82685; NID:g577065; PDB:1A2374.1
C:Genetics:
A:Gene: tXET-B1
C:Superfamily: endoxyloglucan transferase
C:Keywords: glycosidase; hydrolase

Query Match 44.1%; Score 15; DB 2; Length 289;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 219 CWSAASSTSSC 230

RESULT 19

G97525
CG10208 protein (AE003744) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; MUID:21608551; PMID:11743194

A/Accession: G97525
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-318 <KIR>
A/Cross-references: UNIPROT:Q8UFM9, UNIPARC:UPI00000D1B27, GB:AE007869, PIDN:AAK87160.1,
C/Genetics:
A/Gene: AGR_C_2528
A/Map position: circular chromosome

Query Match 44.1%; Score 15; DB 2; Length 318;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 277 CSATHSSDMAAC 288

RESULT 20
T21152
hypothetical protein F20B11.7 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T21152
R/Almscough, R.
submitted to the EMBL Data Library, November 1996
A/Reference number: Z19383
A/Accession: T21152
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-357 <WIL>
A/Cross-references: UNIPROT:Q9YV87, UNIPARC:UPI0000060FAD, EMBL:Z81508, PIDN:CAB04138.1,
A/Experimental source: clone F20B11
C/Genetics:
A/Gene: CBSP:F20B11.7
A/Map position: 5
A/Intons: 52/3; 88/2; 103/3; 154/3; 190/2; 205/3; 241/2; 270/3; 321/3

Query Match 44.1%; Score 15; DB 2; Length 357;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 7 CTATTGTNSAGC 18

RESULT 21
S18765
Sd protein - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C/Accession: S18765
R/Powers, P.A.; Ganetzky, B.
Genetics 129, 133-144, 1991
A/Title: On the components of segregation distortion in Drosophila melanogaster. V. Mole
A/Reference number: S18765; MUID:92038937, PMID:1936954
A/Accession: S18765
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-363 <PCM>
A/Cross-references: UNIPROT:P25722, UNIPARC:UPI000016BD5D, EMBL:X60218, NID:g11012, PIDN
C/Genetics:
A/Gene: FlyBase:sd
A/Cross-references: FlyBase:FBgn0024230

Query Match 44.1%; Score 15; DB 2; Length 363;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 331 CSRTSTFRSSC 342

RESULT 22
T27303
hypothetical protein Y68A4A.10 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T27303
R/Steward, C.
submitted to the EMBL Data Library, January 1998
A/Reference number: Z20340
A/Accession: T27303
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-392 <WIL>
A/Cross-references: UNIPROT:Q9XXP6, UNIPARC:UPI000007BED1, EMBL:AL021503, PIDN:CAA16424.1
A/Experimental source: clone Y68A4A
C/Genetics:
A/Gene: CBSP:Y68A4A.10
A/Map position: 5
A/Intons: 45/2; 74/3; 139/3; 175/2; 190/3; 241/3; 305/3; 356/3

Query Match 44.1%; Score 15; DB 2; Length 392;
Best Local Similarity 16.7%; Pred. No. 1.5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 260 CTATTGMTSTGC 271

RESULT 23
T08928
sucrose cleavage protein homolog T15N24.70 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C/Accession: T08928
R/Beyan, M.; Zimmermann, W.; Grueneisen, A.; Wandt, R.; Bancroft, I.; Mewes, H.W.; Maye
submitted to the Protein Sequence Database, May 1999
A/Reference number: Z16518
A/Accession: T08928
A/Molecule type: DNA
A/Residues: 1-409 <BEV>
A/Cross-references: UNIPROT:Q9UUA2, UNIPARC:UPI000009B309, EMBL:AL078465, GSPDB:GN00062;
A/Experimental source: cultivar Columbia; BMC clone T15N24
C/Genetics:
A/Gene: ATSP:T15N24.70
A/Map position: 4
A/Intons: 126/3; 155/2; 223/1; 299/3

Query Match 44.1%; Score 15; DB 2; Length 409;
Best Local Similarity 16.7%; Pred. No. 1.5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 394 CAASVAVAYTC 405

RESULT 24
S74388
gamma-glutamyl phosphate reductase proA - Synechocystis sp. (strain PCC 6803)
N/Alternate names: hypothetical protein sl10461
C/Species: Synechocystis sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C/Accession: S74388
R/Kaneko, T.; Sato, S.; Kocani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A/Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S74388
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-420 <KAN>
A:Cross-references: UNIPROT:Q55167; UNIPARC:UPI000013948; EMBL:64001; GB:AB001339; NID:96718
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: proA
C:Superfamily: gamma-glutamyl phosphate reductase

Query Match 44.1%; Score 15; DB 2; Length 420;
Best Local Similarity 16.7%; Pred. No. 1.5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15
Db 153 CGASHSHTAATC 164

RESULT 25
D96552
Unknown protein, 74043-75895 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D96552
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizart, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malti, R.; Marzilli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:1130712
A:Accession: D96552
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-423 <STO>
A:Cross-references: UNIPROT:Q9CBJ4; UNIPARC:UPI00000A997F; GB:AE005173; NID:910092345; F
C:Genetics:
A:Gene: FSD21.9
A:Map position: 1
C:Superfamily: Arabidopsis thaliana hypothetical protein F24B22.230

Query Match 44.1%; Score 15; DB 2; Length 423;
Best Local Similarity 16.7%; Pred. No. 1.5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15
Db 306 CTTYTSLKASC 317

RESULT 26
S31126
hypothetical protein F59B2.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S31126
R:Stulson, J.; Du, Z.; Thomas, K.; Wilson, R.; Hillier, L.; Staden, R.; Halloran, N.; Grawkins, T.; Ainscough, R.; Waterston, R.
submitted to the EMBL Data Library, November 1991
A:Description: The C. elegans sequencing project: A beginning.
A:Reference number: S31122
A:Accession: S31126
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-503 <STU>
A:Cross-references: UNIPROT:P34482; UNIPARC:UPI000013B9DD; EMBL:Z11505; NID:96718; PID:9
C:Genetics:
A:Introns: 26/3; 42/2; 121/3; 367/3; 443/1

C:Superfamily: Caenorhabditis elegans hypothetical protein F59B2.6

Query Match 44.1%; Score 15; DB 2; Length 503;
Best Local Similarity 16.7%; Pred. No. 1.5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15
Db 4 CSASTSQLSTFC 15

RESULT 27
B87180
pyruvate (or indolepyruvate) decarboxylase [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 05-Oct-2004
C:Accession: B87180
R:Coile, S.T.; Bigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho R.; Davies, R.M.; Devlin, K.; Duttoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd, eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: B87180
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-569 <STO>
A:Cross-references: UNIPROT:Q9CBDE; UNIPARC:UPI00000CEB3D; GB:AL450380; NID:913093715; P
C:Genetics:
A:Gene: pdc
C:Superfamily: pyruvate decarboxylase/indolepyruvate decarboxylase; thiamin pyrophosphat

Query Match 44.1%; Score 15; DB 2; Length 569;
Best Local Similarity 16.7%; Pred. No. 1.6e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15
Db 140 CSQANLTTATAC 151

RESULT 28
JC7818
metalloprotease I precursor - Alteromonas sp. (Strain O-7)
C:Species: Alteromonas sp. (Strain O-7)
C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 31-Dec-2004
C:Accession: JC7818
R:Myamoto, K.; Teujido, H.; Nukui, E.; Itoh, H.; Kaidzu, Y.; Inamori, Y.
Biosci. Biotechnol. Biochem. 66, 416-421, 2002
A:Title: Isolation and characterization of the genes encoding two metalloproteases (MprI
A:Reference number: JC7818; PMID:1199419; MUID:21994062
A:Accession: JC7818
A:Molecule type: DNA
A:Residues: 1-727 <MTY>
A:Cross-references: UNIPROT:Q8VLQ5; UNIPARC:UPI000017CD49; DDBJ:AB063611
A:Experimental source: strain O-7
C:Comment: This enzyme, an extracellular alkaline metalloprotease, showing high similari
C:Genetics:
A:Gene: mprI
A:Start codon: GTG
C:Superfamily: zinc metalloendopeptidase, neutral protease type (elastase)

Query Match 44.1%; Score 15; DB 2; Length 727;
Best Local Similarity 16.7%; Pred. No. 1.6e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15
Db 581 CNSTSTSNESC 592

RESULT 29

S40382
box A-binding factor - fruit fly (Drosophila melanogaster)
N:Alternate names: ABF; transcription factor dGATAb
C:Species: Drosophila melanogaster
C>Date: 13-Jan-1995 #sequence_revision 06-Sep-1996 #text_change 09-Jul-2004
C/Accession: S40382
R:Abdel, T.; Michelson, A.M.; Maniatis, T.
Development 119, 623-633, 1993
A>Title: A Drosophila GATA family member that binds to Adh regulatory sequences is expressed in the developing eye
A:Reference number: S40382; MUID:94244465; PMID:8187633
A/Accession: S40382
A:Molecule type: mRNA
A:Residues: 1-779 <ABF>
A/Cross-references: UNIPROT:P52172; UNIPARC:UPI0000135F5F; EMBL:X76217; NID:9441491; PID:9441491
C/Comment: This transcriptional activator is the earliest known marker of the developing eye
C/Genetics:
A:Gene: FLYBase:trp
A/Cross-references: FlyBase:FBgn0003507
C/Superfamily: box A-binding factor; GATA-type zinc finger homology
C/Keywords: DNA binding; nucleus; transcription factor; transcription regulation; zinc finger
F:316-369/Domain: GATA-type zinc finger homology <GZF>
F:319-343/Region: zinc finger GATA motif

Query Match 44.1%; Score 15; DB 1; Length 779;
Best Local Similarity 16.7%; Pred. No. 1.7e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 507 CCSISSNNRAC 518

RESULT 30
PYVZAM

spheroidin precursor - Amsacta moorei poxvirus
C/Species: Amsacta moorei poxvirus
C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C/Accession: JQ1436; PQ0289; E41561
R:Banville, M.; Dumas, F.; Triffiro, S.; Arif, B.; Richardson, C.
J. Gen. Virol. 73, 559-566, 1992
A>Title: The predicted amino acid sequence of the spheroidin protein from Amsacta moorei
A:Reference number: JQ1436; MUID:92185464; PMID:1545219
A/Accession: JQ1436
A:Molecule type: DNA
A/Residues: 1-1003 <BAM>
A/Cross-references: UNIPROT:P29815; UNIPARC:UPI00000008A1; GB:M75889; NID:9209629; PIDN:
A/Accession: PQ0289
A:Molecule type: protein
A/Residues: 2-7,'X','9-11,'X',13;86-90,'X',92-98,'X',100-102;532-540;728-750;786-802;810-
A/Cross-references: UNIPARC:UPI00001749E4; UNIPARC:UPI00001749E5; UNIPARC:UPI00001749E6;
R:Hall, R.L.; Moyer, R.W.
J. Virol. 65, 6516-6527, 1991
A>Title: Identification, cloning, and sequencing of a fragment of Amsacta moorei entomop
A:Reference number: A41561; MUID:92046310; PMID:1942245
A/Accession: E41561
A:Molecule type: DNA
A/Residues: 1-1003 <HAL>
A/Cross-references: UNIPARC:UPI00000008A1; GB:M77182; NID:9209631; PIDN:AAA42383.1; PID:
C/Comment: This protein is a major component of the occlusion body which serves to prote
C/Superfamily: Amsacta moorei poxvirus spheroidin
C/Keywords: acetylated amino end; glycoprotein; leucine zipper
F:2-1003/Product: spheroidin #status experimental <MAT>
F:119-140/Region: leucine zipper motif
F:2/Modified site: acetylated amino end (Ser) (in mature form) #status experimental
F:47,109,204,210,259,510,529,547,561,799,901,922,955/Binding site: carboxylate (Asn) (C
F:537,799/Binding site: carboxylate (Asn) (covalent) #status absent

Query Match 44.1%; Score 15; DB 1; Length 1003;
Best Local Similarity 16.7%; Pred. No. 1.8e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

DB 951 CADANSSSDSC 962

RESULT 31

hypochemical protein C34F6.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T19734
R:White, S.
submitted to the EMBL Data Library, November 1996
A/Reference number: Z19171
A/Accession: T19734
A/Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A/Residues: 1-1043 <WIL>
A/Cross-references: UNIPROT:O17644; UNIPARC:UPI00000775AC; EMBL:Z81479; PIDN:CAB03944.1.
A/Experimental source: clone C34F6
C/Genetics:
A:Gene: CESP:C34F6.1
A/Map position: X
A/Intons: 17/3; 62/1; 129/3; 181/1; 292/1; 348/1; 404/1; 665/1; 769/1; 876/1; 977/3

Query Match 44.1%; Score 15; DB 2; Length 1043;
Best Local Similarity 16.7%; Pred. No. 1.8e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 218 CHTGASFTTAC 229

RESULT 32

delta endotoxin - Bacillus thuringiensis
C/Species: Bacillus thuringiensis
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 05-Oct-2004
C/Accession: T18210
R:Narva, K.E.; Payne, J.; Felleison, J.S.
submitted to the EMBL Data Library, August 1994
A>Title: Novel Bacillus thuringiensis toxins active against coleopteran larvae.
A:Reference number: Z18826
A/Accession: T18210
A/Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A/Residues: 1-1166 <NAR>
A/Cross-references: UNIPROT:Q45710; UNIPARC:UPI0000127491; EMBL:U13955; NID:9538377; PID
C/Keywords: delta-endotoxin

Query Match 44.1%; Score 15; DB 2; Length 1166;
Best Local Similarity 16.7%; Pred. No. 1.8e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 466 CTSTANSYKASC 477

RESULT 33

xanthine dehydrogenase homolog T1111.140 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C/Accession: T10236
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Jesse, T.; Heijnen, L.; Vos, P.; Mewes,
submitted to the Protein Sequence Database, June 1999
A/Accession: T10236
A:Molecule type: DNA
A/Residues: 1-1364 <BEV>
A/Cross-references: UNIPROT:Q9SEW45; UNIPARC:UPI000009E008; EMBL:AL079347; GSPDB:GN00062;
A/Experimental source: cultivar Columbia; BAC clone T1111
C/Genetics:

A:Gene: ATSP:T11111.140
A:Map position: 4
A:Introns: 35/1; 66/2; 115/3; 613/3; 694/3; 799/3; 879/3; 974/3; 1048/3; 1102/3; 1147/3;
C:Superfamily: xanthine dehydrogenase; ferredoxin [2Fe-2S] homology
C:Keywords: 2Fe-2S; metalloprotein
F:43,48,51,86/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 44.1%; Score 15; DB 2; Length 1364;
Best Local Similarity 16.7%; Pred. No. 1.9e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 209 CGSKTSEATRC 220

RESULT 34
S65571
Pattern formation protein GNOM - Arabidopsis thaliana

N:Alternate names: EMB30 protein
C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: S65571; S65572

R:Busch, M.; Mayer, U.; Juegens, G.
Mol. Gen. Genet. 250, 681-691, 1996

A:Title: Molecular analysis of the Arabidopsis pattern formation gene GNOM: gene structure
A:Reference number: S65571; MUID:96204508; PMID:8628228
A:Accession: S65571

A:Molecule type: DNA

A:Residues: 1-1451 <BUS>

A:Cross-references: UNIPROT:Q42510; UNIPARC:UPI0000129E82; EMBL:U36433; NID:g1209632; PI
A:Accession: S65572

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-110,'T',112-866,'G',868-1451 <BUS>

A:Cross-references: UNIPARC:UPI000016DA45; EMBL:U36432; NID:g1209630; PID:g1209631

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995
C:Genetics:
A:Gene: GNOM; EMB30
A:Introns: 246/3

Query Match 44.1%; Score 15; DB 2; Length 1451;
Best Local Similarity 16.7%; Pred. No. 1.9e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 22 CTSSNTTTLAC 33

RESULT 35
T25888
hypothetical protein T10H10.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25888

R:Nelson, J.; Langston, Y.
submitted to the EMBL Data Library, December 1996

A:Description: The sequence of C. elegans cosmid T10H10.

A:Reference number: Z20105

A:Accession: T25888

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2098 <NEL>

A:Cross-references: UNIPROT:P91443; UNIPARC:UPI00000756CE; EMBL:U80848; PIDN:AAB37988.1;

A:Experimental source: strain Bristol N2; clone T10H10
C:Genetics:
A:Gene: CESP:T10H10.1
A:Map position: X
A:Introns: 6/3; 43/3; 92/3; 154/2; 226/1; 280/3; 417/2; 472/3; 812/3; 853/3; 887/3; 978/
/3
F:65-720/Domain: myosin motor domain homology <MMO>

Query Match 44.1%; Score 15; DB 2; Length 2098;
Best Local Similarity 16.7%; Pred. No. 2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 1188 CADSATTAELC 1199

RESULT 36

T13806
toucan gene protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13806

R:Grammont, M.; Dastugue, B.; Couderc, J.L.
Development 124, 4917-4926, 1997

A:Title: The Drosophila toucan (tcc) gene is required in germline cells for somatic cell

A:Reference number: Z17769; MUID:98090047; PMID:9362455
A:Accession: T13806

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2176 <GRA>

A:Cross-references: UNIPROT:O46112; UNIPARC:UPI0000083F2C; EMBL:Y14157; NID:g2760521; PI
A:Gene: tcc
A:Cross-references: FlyBase:FBgn0015600

Query Match 44.1%; Score 15; DB 2; Length 2176;
Best Local Similarity 16.7%; Pred. No. 2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 2031 CIRASATSAVC 2042

RESULT 37
S60224
polyketide synthase 1 - Colletotrichum lagenarium

C:Species: Colletotrichum lagenarium

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S60224

R:Takano, Y.; Kubo, Y.; Shimizu, K.; Mise, K.; Okuno, T.; Furusawa, I.
Mol. Gen. Genet. 249, 162-167, 1995

A:Title: Structural analysis of PKS1, a polyketide synthase gene involved in melanin bios

A:Reference number: S60224; MUID:96086932; PMID:7500937
A:Accession: S60224

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-2187 <TAK>

A:Cross-references: UNIPROT:P79068; UNIPARC:UPI000006C308; GB:DB3643; NID:g1208941; PIDN:

A:Gene: PKS1

C:Genetics:
A:Function: melatonin biosynthesis

A:Pathway: melatonin biosynthesis

C:Keywords: carrier protein
F:402-811/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>

F:910-1198/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>

F:166-1736/Domain: acyl carrier protein homology <ACP>

F:1788-1862/Domain: acyl carrier protein homology <ACPI>

Query Match 44.1%; Score 15; DB 2; Length 2187;
Best Local Similarity 16.7%; Pred. No. 2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 553 CSSSAAALNVAC 564

RESULT 38
T28675

alpha-51D immobilization antigen - Paramedium tetraurelia
C:Species: Paramedium tetraurelia
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T28675
R:Schwegmann, K.J.
Submitted to the EMBL Data Library, March 1996
A:Reference number: Z20506
A:Accession: T28675
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2533 <SCH>
A:Cross-references: UNIPROT:P90589, UNIPARC:UPI0000078A73, EMBL:X96400, PIDN:CAA65264.1
C:Genetics:
A:Gene: alpha-51D
A:Genetic code: SGC5
A:Intons: 280/3; 538/2; 1248/2
C:Superfamily: G surface protein

Query Match 44.1%; Score 15; DB 2; Length 2533;
Best Local Similarity 16.7%; Pred. No. 2.1e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 360 CKGTSTTEAAC 371

RESULT 39
T28674
alpha-51D-immobilization antigen - Paramedium tetraurelia
C:Species: Paramedium tetraurelia
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T28674
R:Schmidt, H.J.
submitted to the EMBL Data Library, March 1995
A:Reference number: Z20505
A:Accession: T28674
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2533 <SCH>
A:Cross-references: UNIPROT:Q27183, UNIPARC:UPI0000078B2, EMBL:X85135, NID:g728634, PIDN:
C:Genetics:
A:Genetic code: SGC5
A:Note: alpha-51D
C:Superfamily: G surface protein

Query Match 44.1%; Score 15; DB 2; Length 2533;
Best Local Similarity 16.7%; Pred. No. 2.1e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 360 CKGTSTTEAAC 371

RESULT 40
T31687
surface antigen - Paramedium primaurelia
C:Species: Paramedium primaurelia
C>Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: T31687
R:Bourgain-Guilliet, F.; Caron, F.
Journal of Eukaryot. Microbiol. 43, 303-314, 1996
A:Title: Molecular characterization of the D surface protein gene subfamily in Paramedium
A:Reference number: Z21061; MUID:56313351; PMID:8768434
A:Accession: T31687
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2543 <BOU>
A:Cross-references: UNIPROT:P90649, UNIPARC:UPI0000079A03, EMBL:X96616, NID:g1235576, PIDN:
C:Genetic code: SGC5
A:Superfamily: G surface protein

Query Match 44.1%; Score 15; DB 2; Length 2543;
Best Local Similarity 16.7%; Pred. No. 2.1e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 360 CKGTSTTEAAC 371

RESULT 41
A24420
notch protein - fruit fly (Drosophila melanogaster)
N:Alternate names: neurogenic repetitive locus protein
C:Species: Drosophila melanogaster
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004
C:Accession: A24420; A24768; S09358; A05267
R:Kidd, S.; Kelley, M.R.; Young, M.W.
Mol. Cell. Biol. 6, 3094-3108, 1986
A:Reference number: A24420; MUID:87064624; PMID:3097517
A:Accession: A24420
A:Molecule type: DNA
A:Residues: 1-2703 <KID>
A:Cross-references: UNIPROT:P07207, UNIPARC:UPI000016SC6, GB:K03508, NID:g157991, PIDN:
R:Marion, K.A.; Johansen, K.M.; Xu, T.; Artavanis-Tsakonas, S.
Cell 43, 567-581, 1985
A:Reference number: A24768; MUID:86079539; PMID:3935325
A:Accession: A24768
A:Molecule type: mRNA
A:Residues: 1-48; 'T', 50-118; 'R', 120-230; 'I', 232-256; 'N', 258-266; 'A', 268-872; 'R', 874-958; 'R';
A:Cross-references: UNIPARC:UPI0000173D1F
A:Note: the authors translated the codon ATC for residue 49 as Thr, ATT for residue 2044
R:Tautz, D.
Nucleic Acids Res. 17, 6463-6471, 1989
A:Title: Hypervariability of simple sequences as a general source for polymorphic DNA mar
A:Reference number: S09358; MUID:89385974; PMID:2780284
A:Accession: S09358
A:Molecule type: DNA
A:Residues: 2505-2551, 'OQOO', 2552-2576; 'B', 2578-2604 <TAU>
A:Cross-references: UNIPARC:UPI0000173D20
R:Marion, K.A.; Yedobnick, B.; Flimerty, V.G.; Artavanis-Tsakonas, S.
Cell 40, 55-62, 1985
A:Title: opa: a novel family of transcribed repeats shared by the Notch locus and other c
A:Reference number: A05267; MUID:85099329; PMID:2981631
A:Accession: A05267
A:Molecule type: DNA
A:Residues: 2504-2576; 'E', 2578-2611 <WHA2>
A:Cross-references: UNIPARC:UPI0000173D21
C:Genetics:
A:Gene: notch; opa
A:Cross-references: FlyBase:FBgn0004647
A:Map position: 8.96-9.36
A:Intons: 53/3; 84/3; 171/3; 240/3; 283/3; 2333/3; 2436/3; 2588/3
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology
C:Keywords: differentiation; tandem repeat; transmembrane protein
F:27-43/Domain: transmembrane #status predicted <TMW1>
F:297-328/Domain: EGF homology <EGX1>
F:530-561/Domain: EGF homology <EGF1>
F:558-599/Domain: EGF homology <EGF>
F:968-1019/Domain: EGF homology <EGX2>
F:1064-1095/Domain: EGF homology <EGX3>
F:1187-1218/Domain: EGF homology <EGX3>
F:1746-1765/Domain: transmembrane #status predicted <TMW2>
F:1950-1982/Domain: ankyrin repeat homology <AN1>
F:1983-2015/Domain: ankyrin repeat homology <AN2>
F:1988-2004/Domain: transmembrane #status predicted <TMW3>
F:2017-2049/Domain: ankyrin repeat homology <AN3>
F:2050-2082/Domain: ankyrin repeat homology <AN4>
F:2083-2115/Domain: ankyrin repeat homology <AN5>
F:2538-2568/Region: glutamine-rich
F:2538-2568/Domain: neurogenic repetitive element #status predicted <OPA>

Query Match 44.1%; Score 15; DB 1; Length 2703;

Best Local Similarity 16.7%; Pred. No. 2.1e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 192 CTALAGSSSFTC 203

RESULT 42

hypothetical protein G01D9.5 - Caenorhabditis briggsae

C:Species: Caenorhabditis briggsae

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C:Accession: T11345

R:Waterston, R.

submitted to the EMBL Data Library, April 1996

A:Description: The C. briggsae genome sequencing project.

A:Reference number: 221010

A:Accession: T11345

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4767 <MAT>

A:Cross-references: UNIPROT:Q17301, UNIPARC:UPI000011013B, EMBL:U56248, NID:g1293789, PI

C:Genetics:

A:Insertions: 269/3; 341/3; 853/1; 920/2; 4452/3; 4534/3; 4592/3; 4654/3; 4670/3; 4707/2; 4

A:Note: G01D9.5

C:Keywords: carrier protein; phosphopantetheine; phosphoprotein

F:3472-3541/Domain: acyl carrier protein homology <ACPL>

F:4039-4427/Domain: acetate-CoA ligase homology <ACLI>

F:4447-4514/Domain: acyl carrier protein homology <ACPL>

F:2210.3505/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match

Best Local Similarity 44.1%; Score 15; DB 2; Length 4767;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 1423 CSSSFHALSSAC 1434

RESULT 43

hypothetical protein C41A3.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C:Accession: T15789

R:Bentley, D.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid C41A3.

A:Reference number: Z18404

A:Accession: T15789

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-7829 <BEN>

A:Cross-references: UNIPROT:Q18559; UNIPARC:UPI000011014C; EMBL:U41541; NID:g1109867; PI

C:Genetics:

A:Gene: C41A3.1

A:Insertions: 29/2; 69/3; 141/2; 192/2; 271/3; 321/2; 1000/1; 1098/1; 1127/3; 1168/2; 1282/

2; 7514/3; 7556/3; 7654/3; 7716/3; 7732/3; 7769/2; 7797/3

C:Keywords: carrier protein; phosphopantetheine; phosphoprotein

F:6547-6616/Domain: acyl carrier protein homology <ACP>

F:2832.5271/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match

Best Local Similarity 44.1%; Score 15; DB 2; Length 7829;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 4475 CSSSFYALSSAC 4486

RESULT 44
118346
elastic titin - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004

C:Accession: I18346

R:Labelle, S.; Kolmerer, B.

A:Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.

A:Reference number: A57430; MUID:96026330; PMID:7569978

A:Accession: I18346

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-7962 <RES>

A:Cross-references: UNIPROT:Q10465; UNIPARC:UPI000011010C; EMBL:X90569; NID:g1017426; PI

C:Genetics:

A:Gene: GDB:TTN

A:Cross-references: GDB:127867; OMIM:188840

A:Map position: 2q31-2q31

Query Match
Best Local Similarity 44.1%; Score 15; DB 2; Length 7962;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 381 CTASNEAGSSSC 392

RESULT 45

D22565
R-phycocerythrin beta-2 chain - red alga (Gastrocylindrium coulteri) (fragment)

C:Species: Gastrocylindrium coulteri

C:Date: 07-Mar-1998 #sequence_revision 07-Mar-1998 #text_change 09-Jul-2004

C:Accession: D22565

R:Klotz, A.V.; Glazer, A.N.

J. Biol. Chem. 260, 4856-4863, 1985

A:Title: Characterization of the bilin attachment sites in R-phycocerythrin.

A:Reference number: A22565; MUID:95182601; PMID:3886644

A:Accession: D22565

A:Molecule type: protein

A:Residues: 1-21 <KLO>

A:Cross-references: UNIPROT:Q7M276; UNIPARC:UPI0000174E0D

C:Superfamily: phycocyanin

Query Match
Best Local Similarity 41.2%; Score 14; DB 2; Length 21;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 8 CSMSSSEVASYC 19

RESULT 46

G82754
hypothetical protein XF0861 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C:Accession: G82754

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10510347

A:Note: For a complete list of authors see reference number A59328 below

A:Accession: G82754

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-88 <SIM>

A:Cross-references: UNIPROT:Q9PF17; UNIPARC:UPI00000C254B; GB:AE003925; GB:AE003849; NID:

A:Experimental source: strain 9a5c

R:Simphon, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carraro, D.M.; Carre, H

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincant, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fromt
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B
A;Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsubako, M.H.; Valada, H.; Van Sluys, W.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
A;Genetics:
A;Gene: XF0861

Query Match 41.2%; Score 14; DB 2; Length 88;
Best Local Similarity 16.7%; Pred. No. 0.00052;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15
Db 23 CXXSVTLKSSRC 34

RESULT 47

154781

fibroblast growth factor receptor FGFR-1, secreted splice form - human (fragment)

C;Species: Homo sapiens (man)
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000

C;Accession: 154781

R;Leung, H.Y.; Guillick, W.J.; Lemoine, N.R.

Int. J. Cancer 59, 667-675, 1994

A;Title: Expression and functional activity of fibroblast growth factors and their recep

A;Accession: 154781

A;Reference number: 154781; MUID:95048906; PMID:7960240

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-92 <RES>

A;Cross-references: UNIPARC:UPI0000116DB9; GB:S76658; NID:913557; PIDN:AMD14230.1; PID:

A;Experimental source: pancreatic cancer cell line PT45

C;Genetics:

A;Gene: GDB:FGFR1, FLT2

A;Cross-references: GDB:119913; OMIM:136350

A;Map position: 8p11.2-8p11.1

A;Intons: 58/3

C;Keywords: growth factor receptor

Query Match 41.2%; Score 14; DB 2; Length 92;
Best Local Similarity 16.7%; Pred. No. 0.00053;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15
Db 30 CGAGSASSGRC 41

RESULT 48

T28947

hypothetical protein F07C4.11 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: T28947

R;Miller, N.; Stellyes, L.

submitted to the EMBL Data Library, January 1997

A;Description: The sequence of C. elegans cosmid F07C4.

A;Reference number: Z20546

A;Accession: T28947

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-97 <MTL>

A;Cross-references: UNIPROT:P91216; UNIPARC:UPI000007C37D; EMBL:U80023; PIDN:AAC48019.1;

A;Experimental source: strain Bristol N2; clone F07C4

A;Gene: CESP:F07C4.11

A;Map position: 5

Query Match 41.2%; Score 14; DB 2; Length 97;
Best Local Similarity 16.7%; Pred. No. 0.00053;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15
Db 56 CTTAAADSPSC 67

RESULT 49

C72775

hypothetical protein APE0189 - Aeropyrum pernix (strain KI)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C;Accession: C72775

R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: C72775

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-100 <KAM>

A;Cross-references: UNIPROT:Q9YFR0; UNIPARC:UPI000005DA3D; DDBJ:AP000058; NID:95103388; I

A;Experimental source: strain KI

C;Genetics:

A;Gene: APE0189

C;Superfamily: Aeropyrum pernix hypothetical protein APE0189

Query Match 41.2%; Score 14; DB 2; Length 100;
Best Local Similarity 16.7%; Pred. No. 0.00054;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15
Db 65 CASVSRMARASC 76

RESULT 50

T14780

hypothetical protein DKFZp566B1346.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C;Accession: T14780

R;Ostenmaier, B.; Obermaier, B.; Mewes, H.W.; Gaassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, August 1999

A;Reference number: Z18184

A;Accession: T14780

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-101 <OTT>

A;Cross-references: UNIPARC:UPI000016AC6B; EMBL:AL110253

A;Experimental source: fetal kidney; clone DKFZp566B1346

C;Genetics:

A;Note: DKFZp566B1346.1

Query Match 41.2%; Score 14; DB 2; Length 101;
Best Local Similarity 16.7%; Pred. No. 0.00054;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15
Db 38 CSISVVTQTASC 49

Search completed: January 4, 2006, 16:10:25
Job time : 26.6261 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 ; Search time 117.548 Seconds
(without alignments)
108.037 Million cell updates/sec

Title: US-09-932-322-5
Perfect score: 34
Sequence: 1 XXXXXXXXXXXXXXXX 18

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : Uniprot 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	18	52.9	307 1	SGS3 DROME
2	18	52.9	347 1	OSK340 MOUSE
3	18	52.9	362 2	OSK003 MOUSE
4	18	52.9	367 2	OSK038 ORYZA
5	18	52.9	584 2	OSK043 MOUSE
6	18	52.9	730 2	OSK012 PEA
7	17	50.0	35 2	OSK012 ARATH
8	17	50.0	137 2	OSK060 GIALA
9	17	50.0	137 2	OSK060 GIALA
10	17	50.0	210 2	OSK081 ORYZA
11	17	50.0	234 2	OSK090 GIALA
12	17	50.0	245 2	OSK039 CAEBL
13	17	50.0	306 2	OSK008 CRYNE
14	17	50.0	378 2	OSK002 ORYZA
15	17	50.0	396 2	OSK012 BURK
16	17	50.0	429 2	OSK081 MOUSE
17	17	50.0	540 2	OSK083 STRAW
18	17	50.0	667 2	OSK083 GIALA
19	17	50.0	1031 2	OSK033 EMENT
20	17	50.0	5703 1	MUC5B HUMAN
21	16	47.1	93 2	OSK008 ORYZA
22	16	47.1	100 2	OSK009 ORYZA
23	16	47.1	100 2	OSK009 ORYZA
24	16	47.1	110 1	OSK009 ORYZA
25	16	47.1	144 2	OSK020 SCHUA
26	16	47.1	152 2	OSK009 ABRPE
27	16	47.1	161 2	OSK009 ORYZA
28	16	47.1	172 2	OSK009 ORYZA
29	16	47.1	176 2	OSK009 ORYZA
30	16	47.1	187 2	OSK004 DICDI
31	16	47.1	208 2	OSK004 STRAW

32	16	47.1	210 2	OSK009 ORYZA
33	16	47.1	214 2	OSK009 CRYNE
34	16	47.1	215 2	OSK009 CRYNE
35	16	47.1	215 2	OSK019 CRYNE
36	16	47.1	227 2	OSK009 GLECA
37	16	47.1	230 2	OSK009 GIALA
38	16	47.1	234 2	OSK009 ARATH
39	16	47.1	234 2	OSK009 ARATH
40	16	47.1	244 2	OSK009 ARATH
41	16	47.1	247 2	OSK009 ARATH
42	16	47.1	248 2	OSK009 ARATH
43	16	47.1	257 2	OSK009 ARATH
44	16	47.1	262 2	OSK009 ARATH
45	16	47.1	269 2	OSK009 ARATH
46	16	47.1	278 2	OSK009 ARATH
47	16	47.1	285 2	OSK009 ARATH
48	16	47.1	286 2	OSK009 ARATH
49	16	47.1	287 2	OSK009 ARATH
50	16	47.1	297 2	OSK009 ARATH
51	16	47.1	305 2	OSK009 ARATH
52	16	47.1	309 2	OSK009 ARATH
53	16	47.1	320 2	OSK009 ARATH
54	16	47.1	325 2	OSK009 ARATH
55	16	47.1	329 2	OSK009 ARATH
56	16	47.1	329 2	OSK009 ARATH
57	16	47.1	353 2	OSK009 ARATH
58	16	47.1	366 2	OSK009 ARATH
59	16	47.1	381 2	OSK009 ARATH
60	16	47.1	381 2	OSK009 ARATH
61	16	47.1	426 2	OSK009 ARATH
62	16	47.1	437 2	OSK009 ARATH
63	16	47.1	448 2	OSK009 ARATH
64	16	47.1	465 2	OSK009 ARATH
65	16	47.1	471 2	OSK009 ARATH
66	16	47.1	499 2	OSK009 ARATH
67	16	47.1	504 2	OSK009 ARATH
68	16	47.1	534 2	OSK009 ARATH
69	16	47.1	538 2	OSK009 ARATH
70	16	47.1	539 1	OSK009 ARATH
71	16	47.1	543 2	OSK009 ARATH
72	16	47.1	544 2	OSK009 ARATH
73	16	47.1	556 2	OSK009 ARATH
74	16	47.1	569 2	OSK009 ARATH
75	16	47.1	622 2	OSK009 ARATH
76	16	47.1	667 1	OSK009 ARATH
77	16	47.1	709 2	OSK009 ARATH
78	16	47.1	717 2	OSK009 ARATH
79	16	47.1	719 2	OSK009 ARATH
80	16	47.1	719 2	OSK009 ARATH
81	16	47.1	729 2	OSK009 ARATH
82	16	47.1	760 2	OSK009 ARATH
83	16	47.1	772 2	OSK009 ARATH
84	16	47.1	776 1	OSK009 ARATH
85	16	47.1	776 1	OSK009 ARATH
86	16	47.1	776 1	OSK009 ARATH
87	16	47.1	799 2	OSK009 ARATH
88	16	47.1	868 2	OSK009 ARATH
89	16	47.1	1041 1	OSK009 ARATH
90	16	47.1	1046 2	OSK009 ARATH
91	16	47.1	1089 2	OSK009 ARATH
92	16	47.1	1107 2	OSK009 ARATH
93	16	47.1	1136 2	OSK009 ARATH
94	16	47.1	1167 2	OSK009 ARATH
95	16	47.1	1178 2	OSK009 ARATH
96	16	47.1	1236 2	OSK009 ARATH
97	16	47.1	1341 2	OSK009 ARATH
98	16	47.1	1494 2	OSK009 ARATH
99	16	47.1	1519 2	OSK009 ARATH
100	16	47.1	1592 2	OSK009 ARATH
101	16	47.1	1592 2	OSK009 ARATH
102	16	47.1	35 2	OSK009 ARATH
103	16	47.1	36 2	OSK009 ARATH
104	16	47.1	43 2	OSK009 ARATH

981	14	41.2	309	2	08P004_XMNAc	08P054_xanthomonas
982	14	41.2	311	2	08Y6XZ_VYIRU	08Y6XZ_iris yellow
983	14	41.2	311	2	091PB3_VYIRU	091PB3_iris yellow
984	14	41.2	312	2	016417_CAEEL	016417_caeornithabdi
985	14	41.2	315	2	Q7ONT6_GIALA	Q7ONT6_giardia lam
986	14	41.2	316	2	0680KZ_ARATH	0680KZ_arabidopsis
987	14	41.2	316	2	08L7U1_ARATH	08L7U1_arabidopsis
988	14	41.2	316	2	08L8A1_ARATH	08L8A1_arabidopsis
989	14	41.2	317	2	06FHE1_HOMAN	06FHE1_homo sapiens
990	14	41.2	317	2	08WVZ0_TRINI	08WVZ0_trichoplus
991	14	41.2	317	2	08BHT0_MOUSE	08BHT0_mus musculus
992	14	41.2	318	2	06SANT_9SAUR	06SANT_rhineci s
993	14	41.2	321	1	HEMH_WIGSR	08d226_wigglesworth
994	14	41.2	321	2	Q7RI09_GIALA	Q7RI09_giardia lam
995	14	41.2	321	2	06NGE1_CORDI	06NGE1_corynebacte
996	14	41.2	321	2	06S9M6_9SAUR	06S9M6_xenochrophi
997	14	41.2	321	2	06S9P5_9SAUR	06S9P5_hemeterophi
998	14	41.2	321	2	06S9A7_9SAUR	06S9A7_platyceps r
999	14	41.2	321	2	06S9A2_9SAUR	06S9A2_platyceps r
1000	14	41.2	321	2	06S9A3_9SAUR	06S9A3_platyceps k

ALIGNMENTS

RESULT 1	SGS3_DROME	STANDARD;	PRT;	307 AA.
ID	SGS3_DROME	STANDARD;	PRT;	307 AA.
AC	P02840; Q9VU72;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Salivary glue protein Sgs-3 precursor.			
GN	Name=SGS3; ORFNames=CG11702.			
OS	Drosophila melanogaster (fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_Taxid=7227;			
NP	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=83294545; PubMed=6411930;			
RA	Garfinkel M.D., Pruitt R.B., Meyerowitz E.M.;			
RT	"DNA sequences, gene regulation and modular protein evolution in the			
RT	Drosophila 68C glue gene cluster.;"			
RL	J. Mol. Biol. 168:765-789(1993).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RC	STRAIN=Berkeley;			
RX	MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;			
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amannikides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abblitt J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,			
RA	Ballew R.M., Baugh A., Baxendale J., Bayraktaroglu U., Beasley E.M.,			
RA	Beeson K.Y., Beres P.V., Bernan B.P., Bhandari D., Bolshakov S.,			
RA	Botkova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cickstein E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlie C., Davenport L.B., Davies P.,			
RA	de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,			
RA	Fowler C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glaser K.,			
RA	Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jaitai M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Kelchum K.A.,			
RA	Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Lai X., Lai Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,			
RA	Liou X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			

RA Mekhlouov G., Mikhina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Hoy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacle D.,
RA Palazolo M., Pltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye Y., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
[3]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RX Mista S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminck J.S., Millburn G.H., Prochack S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu I., Bertram B.P.,
RA Bettencourt B.R., Celniker S.B., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter U., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
[4]
RP NUCLEOTIDE SEQUENCE OF 1-28.
RX MEDLINE=88332966; PubMed=3138416;
RA Martin C.H., Mayeda C.A., Meyerowitz E.M.;
RT "Evolution and expression of the Sgs-3 glue gene of *Drosophila*.";
RL J. Mol. Biol. 201:273-287(1988).
[5]
RP DEVELOPMENTAL STAGE.
RX MEDLINE=94038699; PubMed=8223281;
RA Hueb F., Ruiz C., Richards G.;
RT "Puffs and PCR: the in vivo dynamics of early gene expression during
RT ecdysose responses in *Drosophila*.";
RL Development 118:613-627(1993).
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: Salivary gland specific.
CC -I- DEVELOPMENTAL STAGE: In the salivary glands of mid instar larvae
CC levels dramatically increase during puff stage 1 at 98-106 hours
CC of development. Levels remain constant and abundant in late larvae
CC until puff stage 10, then decrease by stage 11.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation --
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

DR EMBL; X01918; CAA25994.1; -; Genomic DNA.
DR EMBL; AE003544; AAF50056.1; -; Genomic DNA.
DR EMBL; X78932; CAA55154.1; -; Genomic DNA.
DR PIR; A03329; GSFF3.
DR Ensemble; CG11720; *Drosophila melanogaster*.
DR FlyBase; FBgn0003373; Sgs3.
KW Repeat; signal.
KW SIGNAL.
FT CHAIN 1 23 potential,
FT 24 307 Salivary glue protein Sgs-3.
FT SEQUENCE 307 AA; 32196 MW; 45803DEB16C418BC CRC64;

Query Match 52.9%; Score 18; DB 1; Length 307;
Best Local Similarity 16.7%; Pred. No. 4,9e-11;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
4 CXXXXXXXXXC 15
44 CTTTTTTTTTC 55
DB

```

RESULT 2
ORF340 MOUSE
ID ORF340_MOUSE PRELIMINARY; PRT; 347 AA.
AC ORF340;
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
DE 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE 463241106Rik protein (Fragment).
GN Name=463241106Rik;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RC TISSUE=Mammary tumor metastasized to lung. MMTV-LTR/Mnt1 model.
RX Expression driven by an MMTV-LTR enhancer;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullanb S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RC TISSUE=Mammary tumor metastasized to lung. MMTV-LTR/Mnt1 model.
RC Expression driven by an MMTV-LTR enhancer;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028854; AAH28854.1; -; mRNA.
DR MGI; MGI:1925998; 463241106Rik.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001159; RA.
DR PROSITE; PS50200; RA; 1.
DT NON TR
FT 1
SQ SEQUENCE 347 AA; 39161 MW; 306B40D38A14B3DE CRC64;
Query Match 52.9%; Score 18; DB 2; Length 347;
Best Local Similarity 16.7%; Freq. NO. 5.1e-11;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Cy 4 CXXXXXXXXXXC 15
Db 135 CSSTSSSTASCC 146

```

```

DE enriched library, clone:6030474M07 product:hypothetical Serine-rich
DE region/Aminoacyl-transfer RNA synthetases class-II containing protein.
DE full insert sequence.
GN Name=463241106Rik;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi U., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
RA Saito T., Okazaki Y., Gojobori T., Bonb H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuhl P., Lewis S., Matsuo Y., Nikaido I., Peesle G., Quackenbush J.,
RA Schirral L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Barsh G.,
RA Blake J., Botfield D., Boujuna N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guenrich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mommaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seye T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitlatch C., Wilting L.,
RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA The PANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitunui T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto K., Matsuno K., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.

```

RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Pakuda S., Furuno M., Hanagata T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai K., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Yamatsu M., Hayashizaki Y.,
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK031660; BAC27498.1; -; mRNA.
 DR MGI; MGI:192598; 463241J06Rik.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000159; RA.
 DR SMART; SM00314; RA; 1.
 DR PROSITE; PS50200; RA; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 362 AA; 40934 MW; 32B1047AC07B987 CRC64;

Query Match 52.9%; Score 18; DB 2; Length 362;
 Best Local Similarity 16.7%; Pred. No. 5.2e-11;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15
 Db 199 CSSTSSSTASC 210

RESULT 4
 ID 06AV38_ORYSA PRELIMINARY; PRT; 367 AA.
 AC 06AV38;
 DT 25-OCT-2004 (TREMblrel. 28, Created)
 DT 25-OCT-2004 (TREMblrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMblrel. 28, Last annotation update)
 DE Putative Cyclin.
 GN Name=OSUNBa0063J18.4;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatocidae; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
 RA Overton II L.L., Teltzin T., Kim M.M., Bera J.J., Jin S.S.,
 RA Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsieh J., Blunt S.,
 RA Vanaken S.S., Riedmiller S.B., Uteback T.T., Feldblyum T.V.,
 RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
 RA White O., Salzberg S.L., Frazer C.M.;
 RT "Oryza sativa chromosome 3 BAC OSUNBa0063J18 genomic sequence."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell R.;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 CC -; SIMILARITY: Belongs to the cyclin family.
 DR EMBL; AC107206; AAT77041.1; -; Genomic_DNA.
 DR Gramene; 06AV38; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0000074; P:regulation of cell cycle; IEA.
 DR InterPro; IPR006670; Cyclin.
 DR InterPro; IPR004367; Cyclin_Cterm.
 DR InterPro; IPR006671; Cyclin_N.
 DR Pfam; PF02984; Cyclin_C; 1.
 DR Pfam; PF00134; Cyclin_N; 1.
 DR SMART; SM00385; CYCLIN; 1.
 KM Cyclin.
 SQ SEQUENCE 367 AA; 38727 MW; 023BF8C0CDAD1D79 CRC64;

Query Match 52.9%; Score 18; DB 2; Length 362;
 Best Local Similarity 16.7%; Pred. No. 5.2e-11;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15
 Db 75 CSSTSSSTSSSC 86

RESULT 5
 ID 08BL43_MOUSE PRELIMINARY; PRT; 584 AA.
 AC 08BL43;
 DT 01-MAR-2003 (TREMblrel. 23, Created)
 DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
 DE Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length
 DE enriched library, clone:B230384H21 product:hypothetical Serine-rich
 DE region/Aminoacyl-transfer RNA synthetases class-II containing protein,
 DE full insert sequence. (Fragment).
 GN Name=463241J06Rik;
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi Y., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nakai I., Pesole G., Quackenbush J.,
 RA Schirai L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barb G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bona M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kontecki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Suganara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).
 [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CS7BL/6J; TISSUE=Corpora quadrigemina;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600.
 RA Shibata K., Itoh M., Atzawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Komu H., Akiyama J., Nishi K., Kitsuami T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasahigaki K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CS7BL/6J; TISSUE=Corpora quadrigemina;
 RA Adachi S., Atzawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hoti F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Komu H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK046433; BAC32726.1; - mRNA.
 DR MGI: MGI:1925998; 463241J06R1X.
 DR GO: GO:0007165; P: signal transduction; IEA.
 DR InterPro: IPR001357; RA.
 DR SMART: SM00314; RA; 1.
 DR PROSITE: PS50200; RA; 1.
 KM Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 584 AA; 65345 MW; 941D03439114C6E CRC64;
 Query Match 52.9%; Score 18; DB 2; Length 584;
 Best Local Similarity 16.7%; Pred. No. 5.9e-11;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 Qy 4 CXXXXXXXXXXC 15
 Db 275 CSSTSSSTASC 286
 RESULT 6
 OBLP13_PEA
 ID OBLP13_PEA PRELIMINARY; PRT; 730 AA.
 AC OBLP13;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE ABA insensitive 3.
 GN Name=ab13;
 OS Pisum sativum (garden pea).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
 OC NCBI_TaxID=3888;
 OX NCBI_TaxID=3888;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Axillary bud;
 RA Nakako A., Mori H.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB080195; BAC10553.1; - mRNA.
 DR GO: GO:0003677; F: DNA binding; IEA.
 DR GO: GO:0006355; P: regulation of transcription, DNA-dependent; IEA.
 DR InterPro: IPR003340; TF_B3.
 DR Pfam: PF02362; B3; 1.

DR PROSITE: PS50863; B3; 1.
 SQ SEQUENCE 730 AA; 80868 MW; 0A5BE9A7350AB8B2 CRC64;
 Query Match 52.9%; Score 18; DB 2; Length 730;
 Best Local Similarity 16.7%; Pred. No. 6.2e-11;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 Qy 4 CXXXXXXXXXXC 15
 Db 55 CSSSSASSSSC 66
 RESULT 7
 O9LQ12_ARATH
 ID O9LQ12_ARATH PRELIMINARY; PRT; 35 AA.
 AC O9LQ12;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE F1504.2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Ecker J.R.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Chouk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafi H., Bel B., Chin C., Chiu J., Choi E.,
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsy N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Toriumi M., Vayberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC007887; AAF79343.1; - genomic DNA.
 SQ SEQUENCE 35 AA; 3663 MW; CB4370A9D6A5D1F4 CRC64;
 Query Match 50.0%; Score 17; DB 2; Length 35;
 Best Local Similarity 16.7%; Pred. No. 1.9e-09;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 Qy 4 CXXXXXXXXXXC 15
 Db 15 CSIASTAATSC 26
 RESULT 8
 Q24960_GIALA
 ID Q24960_GIALA PRELIMINARY; PRT; 137 AA.
 AC Q24960;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Tirophozolite surface protein 11 (fragment).
 GN Name=TSPI1/L2;
 OS Giardia lamblia (Giardia intestinalis).
 CC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
 OC NCBI_TaxID=5741;
 OX NCBI_TaxID=5741;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC MEDLINE=98096860; PubMed=9435134;
 RA Ey P.L., Mansouri M., Kulda J., Nohynkova E., Monis P.T.,
 RA Andrews R.H., Mayrhofer G.;
 RT "Genetic analysis of Giardia from hoofed farm animals reveals
 aridocacyl-specific and potentially zoonotic genotypes.";
 RL J. Eukaryot. Microbiol. 44:626-635(1997).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.

RA Ey P.L., Mansouri M., Kulda J., Nohynkova E., Monis P.T.,
 Andrews R.H., Mayrhofer G.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U47631; AAB05399.1; -, Genomic_DNA.
 DR InterPro; IPR006209; EGF_1like.
 DR InterPro; IPR005127; Giardia_VSP.
 DR Pfam; PF03302; VSP; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 FT NON_TER 1
 FT NON_TER 137
 SQ SEQUENCE 137 AA; 13875 MW; 4234696B5ECF37D CRC64;
 Query Match 50.0%; Score 17; DB 2; Length 137;
 Best Local Similarity 16.7%; Pred. No. 2.7e-09;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 4 CXXXXXXXXXXC 15
 DB 22 CASATARAATC 33

RESULT 9
 ID Q24987_GIALA PRELIMINARY; PRT; 137 AA.
 AC Q24987;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Surface protein (Fragment).
 GN Name=VSP;
 OS Giardia lamblia (Giardia intestinalis).
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
 OX NCBI_TaxId=5741;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Ey P.L., Darby J.M., Mayrhofer G.;
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL; L16973; AAB59198.1; -, Genomic_DNA.
 DR InterPro; IPR005127; Giardia_VSP.
 DR Pfam; PF03302; VSP; 1.
 FT NON_TER 1
 FT NON_TER 137
 SQ SEQUENCE 137 AA; 14148 MW; A340072153005D58 CRC64;
 Query Match 50.0%; Score 17; DB 2; Length 137;
 Best Local Similarity 16.7%; Pred. No. 2.7e-09;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 4 CXXXXXXXXXXC 15
 DB 22 CTSTTARTATC 33

RESULT 10
 ID Q7XHS1_ORYSA PRELIMINARY; PRT; 210 AA.
 AC Q7XHS1;
 DT 01-OCT-2003 (TReMBLrel. 25, Created)
 DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE 2Fe-2S Iron-sulfur cluster protein-like.
 GN Name=P0477A12.11;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxId=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare (Gh3) genomic DNA, chromosome 7, PAC
 clone:P0477A12.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP005190; BAC80058.1; -, Genomic_DNA.
 DR Gramene; Q7XHS1: -;
 DR GO; GO:0005489; F:electron transporter activity; IEA.
 DR GO; GO:0005506; F:iron ion binding; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR01041; Ferredoxin.
 DR InterPro; IPR012675; Ferredoxin_fold.
 DR Pfam; PF00111; Fer2; 1
 SQ SEQUENCE 210 AA; 22629 MW; 6BCDC4F408B21E1E CRC64;
 Query Match 50.0%; Score 17; DB 2; Length 210;
 Best Local Similarity 16.7%; Pred. No. 3e-09;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 4 CXXXXXXXXXXC 15
 DB 25 CAVATATTSSC 36

RESULT 11
 ID Q9XY90_GIALA PRELIMINARY; PRT; 234 AA.
 AC Q9XY90;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Variant-specific surface protein (Fragment).
 GN Name=vsp417-4;
 OS Giardia lamblia (Giardia intestinalis).
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
 OX NCBI_TaxId=5741;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AQ-2;
 RA MEDLINE=99229863; PubMed=10215024; DOI=10.1016/S0166-6851(98)00183-2;
 RA Ey P.L., Darby J.M., Mayrhofer G.;
 RT "A new locus (vsp417-4) belonging to the tsa417-like subfamily of
 RT variant-specific surface protein genes in Giardia intestinalis.";
 RL Mol. Biochem. Parasitol. 99:55-68(1999).
 DR EMBL; AF065600; AAD28789.1; -, Genomic_DNA.
 DR InterPro; IPR006212; Furin_repeat.
 DR InterPro; IPR005127; Giardia_VSP.
 DR InterPro; IPR006210; IEGF.
 DR Pfam; PF03302; VSP; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00261; FU; 1.
 FT NON_TER 1
 FT NON_TER 234
 SQ SEQUENCE 234 AA; 23565 MW; FDC75E280AF7D517 CRC64;
 Query Match 50.0%; Score 17; DB 2; Length 234;
 Best Local Similarity 16.7%; Pred. No. 3.1e-09;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 4 CXXXXXXXXXXC 15
 DB 22 CTGATARAATC 33

RESULT 12
 ID O18039_CABEL PRELIMINARY; PRT; 245 AA.
 AC O18039;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Hypothetical protein T06C12.14.
 GN ORNames=T06C12.14;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderae; Caenorhabditis.
 OX NCBI_TaxId=6239;

RM NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9551916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z81116; CAB03300.1; -; Genomic_DNA.
DR PIR; T24565; T24565.
DR Ensemble; T06C12.14; Caenorhabditis elegans.
DR WormBase; WBGene00011521; T06C12.14.
DR WormPeP; T06C12.14; CE16364.
DR InterPro; IPR003582; SHKT.
DR Pfam; PF01549; SHKT; 3.
DR SMART; SM00254; SHKT; 4.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 245 AA; 25713 MW; BDE40649C966769A CRC64;

Query Match 50.0%; Score 17; DB 2; Length 245;
Best Local Similarity 16.7%; Pred. No. 3.1e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 193 CASATTRRSSTC 204

RESULT 13
Q5WQW8 CRYNE
ID Q5WQW8 CRYNE PRELIMINARY; PRT; 306 AA.
AC Q5WQW8;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Sec53p.
GN Name=SEC53;
OS Cryptococcus neoformans var. neoformans.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=40410;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jambon G.;
RT "SEC53 encodes a putative phosphomannomutase.";
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY426175; AAR84595.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004615; F:phosphomannomutase activity; IEA.
DR GO; GO:0019307; F:mannose biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR006379; HAD_SF_IIB.
DR InterPro; IPR005002; PWW_M.
DR Pfam; PF03332; PWW; 1.
DR TIGRFAMs; TIGR01484; HAD-SF-IIB; 1.
SQ SEQUENCE 306 AA; 33844 MW; FB9DB2CFLA3EDDEA CRC64;

Query Match 50.0%; Score 17; DB 2; Length 306;
Best Local Similarity 16.7%; Pred. No. 3.3e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 5 CSSATTIATTC 16

RESULT 14
Q6ASQ2 ORYSA
ID Q6ASQ2 ORYSA PRELIMINARY; PRT; 378 AA.
AC Q6ASQ2;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)

DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein OSUNBA0077322.21.
GN Name=OSUNBA0077322.21;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chow T.-Y., Hsing Y.-I. C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Lwu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.,
RT "Oryza sativa BAC OSUNBA0077322 genomic sequence.";
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC148814; AAT93980.1; -; Genomic_DNA.
DR Gramene; Q6ASQ2; -;
KW Hypothetical protein.
SQ SEQUENCE 378 AA; 37703 MW; 342586481D07BFB5B CRC64;

Query Match 50.0%; Score 17; DB 2; Length 378;
Best Local Similarity 16.7%; Pred. No. 3.5e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 180 CVAAAAAATTAAC 191

RESULT 15
Q4LGT2_9BURK
ID Q4LGT2_9BURK PRELIMINARY; PRT; 396 AA.
AC Q4LGT2;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Bcen2424DRAFT_0168;
OS Burkholderia cenocepacia H12424.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cepacia complex.
OX NCBI_TaxID=331272;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA US DOE Joint Genome Institute (JGI-ORNL);
RC STRAIN=H12424;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Igarashi S., Pittluck S., Richardson P.,
RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia H12424.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H12424;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Laitimer F., Land M.;
RT "Annotation of the draft genome assembly of Burkholderia cenocepacia H12424.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AALH01000139; EAM15270.1; -; Genomic_DNA.
DR Hypothetical protein.
SQ SEQUENCE 396 AA; 41189 MW; C829063DC312B057 CRC64;

Query Match 50.0%; Score 17; DB 2; Length 396;
Best Local Similarity 16.7%; Pred. No. 3.5e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXX 15
 Db 18 CSSSDAAAAAAC 29

RESULT 16
 ID Q8C8X1_MOUSE PRELIMINARY; PRT; 429 AA.
 Q8C8X1;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mus musculus adult retina cDNA, RIKEN full-length enriched library,
 DE clone: A930006D11 product: hypothetical protein, full insert
 DE sequence.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CS7BL/6J; TISSUE=Retina;
 RA MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CS7BL/6J; TISSUE=Retina;
 RA MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
 RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Futuno M., Aono H., Baldarelli R., Barri G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohleutski S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CS7BL/6J; TISSUE=Retina;
 RA The FANTOM Consortium,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CS7BL/6J; TISSUE=Retina;
 RA MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CS7BL/6J; TISSUE=Retina;
 RA MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sakaki N., Carninci P.,

PA Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
 PA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 PA Yamamoto R., Matsumoto H., Sakaguchi S., Ilegami T., Kashiwagi K.,
 PA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Katsuhiki M.,
 PA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 PA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CS7BL/6J; TISSUE=Retina;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Hayashida S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohnato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Saeki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akaiura S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Yamamoto M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK044299; BAC31860.1; -, mRNA.
 DR Ensembl; ENSMUSG00000014198; Mus musculus.
 DR GO; GO:0005634; C:nucleus; Mus.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR InterPro; IPR007087; ZnF_C2H2.
 DR InterPro; IPR003604; ZnF_U1.
 DR SMART; SM00355; ZnF_C2H2; 3.
 DR SMART; SM00451; ZnF_U1; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_3.
 KM Hypothetical protein.
 SQ SEQUENCE 429 AA; 44564 MW; 211C9F431BF7FEF CRC64;

Query Match 50.0%; Score 17; DB 2; Length 429;
 Best Local Similarity 16.7%; Pred. No. 3.6e-09;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXX 15
 Db 181 CDAAASSSSSSC 192

RESULT 17
 ID Q93H33_STRAW PRELIMINARY; PRT; 540 AA.
 Q93H33;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 13-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Transport system protein (putative oligopeptide ABC transporter
 DE substrate-binding protein).
 GN Name=OPB2; OrderedLocustNames=SAV3149;
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomycetes.
 NCBI_TaxID=33903;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RA MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinoue M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hatтори M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: deducing the ability of producing secondary
 RT metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbr820;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sasaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism *Streptomyces avermitilis*,"
 RL Nat. Biotechnol. 21:526-531 (2003).
 DR EMBL: AB070952; BAB69358.1; -; Genomic DNA.
 DR GO: GO:0005215; P:transporter activity; IEA.
 DR GO: GO:0006810; P:transporter activity; IEA.
 DR InterPro: IPR000914; SBP_bac_5.
 DR Pfam: PF00496; SBP_bac_5; 1.
 KM Complete proteome.
 SQ SEQUENCE 540 AA; 58574 MW; 0052FEAD151B442C CRC64;

Query Match 50.0%; Score 17; DB 2; Length 540;
 Best Local Similarity 16.7%; Pred. No. 3.8e-09;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
 Db 12 CAAAAAATATC 23

RESULT 18
 ID 09XTK3_GIALA PRELIMINARY; PRT; 667 AA.
 AC 09XTK3;
 DT 01-NOV-1999 (TREMBLrel. 12; Created)
 DT 01-NOV-1999 (TREMBLrel. 12; Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26; Last annotation update)
 DE VS417-3/A-II.
 OS Name=vsp417-3/A-II;
 GN Giardia lamblia (Giardia intestinalis).
 CC Eukaryota; Diplomonadida; Hexamitidae; Giardiae; Giardia.
 OC NCBI_TaxID=5741;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=BRIS/83/HEPU/136;
 RA MEDLINE=99053029; PubMed=9836309;
 RA By P.L., Darby J.M., Mayrhofer G.;
 RT "Comparison of tsaa17-like variant surface protein (VSP)
 RT genes in Giardia intestinalis and identification of a novel locus in
 RT genetic group II isolates,"
 RL Parasitology 117:445-455 (1998).
 DR EMBL: AF033584; AAD03497.1; -; Genomic DNA.
 DR HSP: O16119; IEZG.
 DR InterPro: IPR006212; Furin_repeat.
 DR InterPro: IPR005127; Giardia_VSP.
 DR InterPro: IPR006210; IEGF.
 DR Pfam: PF03302; VSP; 1.
 DR SMART: SMO0181; EGF; 2.
 DR SMART: SMO0261; FU; 4.
 SQ SEQUENCE 667 AA; 69123 MW; 77C64CF59441C0C CRC64;

Query Match 50.0%; Score 17; DB 2; Length 667;
 Best Local Similarity 16.7%; Pred. No. 4e-09;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
 Db 448 CTSTTATATC 459

RESULT 19
 ID 05B3H3_EMENTI PRELIMINARY; PRT; 1031 AA.
 AC 05B3H3;
 DT 10-MAY-2005 (TREMBLrel. 30; Created)
 DT 10-MAY-2005 (TREMBLrel. 30; Last sequence update)
 DT 10-MAY-2005 (TREMBLrel. 30; Last annotation update)
 DE Hypothetical protein.

GN ORFNames=AM4907.2;
 OS Aspergillus nidulans FGSC A4.
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 CC Eurotiiales; Trichocomaceae; Emicellae.
 OC NCBI_TaxID=227321;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=FGSC A4;
 RA Birren B., Nisbun C., Abouelleil A., Allen N., Anderson S.,
 RA Archen H.M., Barna N., Bastien V., Bloom T., Boguslavsky L.,
 RA Boukhvalter B., Butler J., Calvo S.E., Camarata J., Chang Y.,
 RA Choepel Y., Collamore A., Cook A., Cooke P., Corum B., Deatellano K.,
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
 RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
 RA Gardina S., Gierre S., Graham L., Grand-Pierre N., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Horon L., Hulme W., Iliev I.,
 RA Kelle D., Johnson R., Jones C., Kamel A., Karatae A.,
 RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
 RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
 RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Menus L.,
 RA Mihova T., Mienga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
 RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
 RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
 RA Rachupka A., Ramasamy U., Raymond C., Reta R., Rise C., Rogov P.,
 RA Roman J., Schauer S., Schupack R., Seaman S., Severy P., Smirnov S.,
 RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
 RA Talamas J., Testaye S., Theodore J., Topham K., Travers M.,
 RA Vasilev H., Venkatarman V.S., Viel R., Vo A., Wang S., Wilson B.,
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
 RA Lander E.;
 RT "Genome Sequence of *Aspergillus nidulans*,"
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL: AACD0100084; EAA60985.1; -; Genomic DNA.
 KM Hypothetical protein.
 SQ SEQUENCE 1031 AA; 113626 MW; 6616443769CF2EF CRC64;

Query Match 50.0%; Score 17; DB 2; Length 1031;
 Best Local Similarity 16.7%; Pred. No. 4.5e-09;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
 Db 743 CSTSSSASASTC 754

RESULT 20
 ID MUC5B_HUMAN STANDARD; PRT; 5703 AA.
 AC 09HC84; 000447; 000573; 014985; 015494; 095291; 095451; Q14881;
 AC 099552; Q9UE28;
 DT 28-FEB-2003 (Ref. 41; Created)
 DT 28-FEB-2003 (Ref. 41; Last sequence update)
 DT 13-SEP-2005 (Ref. 48; Last annotation update)
 DE Mucin-5B precursor (Mucin 5 subtype B, tracheobronchial) (High
 DE molecular weight salivary mucin M61) (bublingual gland mucin).
 GN Name=MUC5B; Synonyms=MUC5;
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE OF 1-1594.
 RA Chen Y., Di Y.P., Wu R.;
 RT "Molecular cloning of the amino-terminal and 5'-flanking region of the
 RT human MUC5B mucin gene,"
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE OF 1-1325.
 RX MEDLINE=9909274; PubMed=9790959; DOI=10.1006/dbrc.1998.9469;

RA Offner G.D., Nunes D.P., Keates A.C., Afchal N.H., Troxler R.F.;
 RT "The amino-terminal sequence of MUC5B contains conserved
 RT multifunctional D domains: implications for tissue-specific mucin
 RT functions."; *Biochem. Res. Commun.* 251:350-355(1998).
 RL Biochem. Res. Commun. 251:350-355(1998).
 RN [3]
 RN NUCLEOTIDE SEQUENCE OF 40-1324.
 RX MEDLINE=99029323; PubMed=9804771; DOI=10.1074/jbc.273.46.30157;
 RA Desseyn J.-L., Buisine M.P., Porchet N., Aubert J.-P., Laine A.;
 RT "Genomic organization of the human mucin gene MUC5B: cDNA and genomic
 RT sequences upstream of the large central exon."; *J. Biol. Chem.* 273:30157-30164(1998).
 RL J. Biol. Chem. 273:30157-30164(1998).
 RN [4]
 RN NUCLEOTIDE SEQUENCE OF 1326-4895.
 RC TISSUE=Placenta;
 RX MEDLINE=97166151; PubMed=9013550; DOI=10.1074/jbc.272.6.3168;
 RA Desseyn J.-L., Guyonnet-Duperrat V., Porchet N., Aubert J.-P.,
 RA Laine A.;
 RT "Human mucin gene MUC5B, the 10.7 kb large central exon encodes
 RT various alternate subdomains resulting in a super-repeat. Structural
 RT evidence for a 11p15.5 gene family."; *J. Biol. Chem.* 272:3168-3178(1997).
 RL J. Biol. Chem. 272:3168-3178(1997).
 RN [5]
 RN NUCLEOTIDE SEQUENCE OF 4057-4480.
 RC TISSUE=Salivary gland;
 RX MEDLINE=97292540; PubMed=9147051;
 RA Nielsen P.A., Bennett E.P., Wandall H.H., Therkildsen M.H.,
 RA Hennibal J., Clausen H.;
 RT "Identification of a major human high molecular weight salivary mucin
 RT (MG1) as tracheobronchial mucin MUC5B."; *J. Biol. Chem.* 272:413-419(1997).
 RL J. Biol. Chem. 272:413-419(1997).
 RN [6]
 RN NUCLEOTIDE SEQUENCE OF 4721-5703.
 RC TISSUE=Gall bladder;
 RX MEDLINE=97293228; PubMed=9164870;
 RA Keates A.C., Nunes D.P., Afchal N.H., Troxler R.F., Offner G.D.;
 RT "Molecular cloning of a major human gall bladder mucin: complete C-
 RT terminal sequence and genomic organization of MUC5B."; *Biochem. J.* 324:295-303(1997).
 RL Biochem. J. 324:295-303(1997).
 RN [7]
 RN NUCLEOTIDE SEQUENCE OF 4809-5687.
 RC TISSUE=Sublingual gland;
 RX MEDLINE=96123551; PubMed=8554565;
 RA Troxler R.F., Offner G.D., Zhang F., Ioncheva I., Oppenheim F.G.;
 RT "Molecular cloning of a novel high molecular weight mucin (MG1) from
 RT human sublingual gland."; *Biochem. Res. Commun.* 217:1112-1119(1995).
 RL Biochem. Res. Commun. 217:1112-1119(1995).
 RN [8]
 RN NUCLEOTIDE SEQUENCE OF 4859-5703.
 RC TISSUE=Placenta;
 RX MEDLINE=97347489; PubMed=9201995; DOI=10.1074/jbc.272.27.16873;
 RA Desseyn J.-L., Aubert J.-P., Porchet N., Laine A.;
 RT "Genomic organization of the 3 region of the human MUC5B mucin."; *J. Biol. Chem.* 272:16873-16883(1997).
 RL J. Biol. Chem. 272:16873-16883(1997).
 CC - FUNCTION: Salivary mucin that is thought to contribute to the
 CC lubricating and viscoelastic properties of whole saliva.
 CC - SUBCELLULAR LOCATION: Secreted.
 CC - TISSUE SPECIFICITY: Expressed mainly in bronchus glands and also
 CC in submaxillary glands, endocervix, gall bladder, and pancreas.
 CC - PTM: Highly glycosylated.
 CC - SIMILARITY: Contains 1 CTCK (C-terminal cysteine knot-like) domain.
 CC - SIMILARITY: Contains 1 TIL (trypsin inhibitory-like) domain.
 CC - SIMILARITY: Contains 3 WFMC domains.
 CC - SIMILARITY: Contains 4 WFMD domains.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC DR EMBL; AF107890; AAC3673.1; -; Genomic DNA.
 CC DR EMBL; AF086604; AAC67545.1; -; mRNA.

DR EMBL; AJ004862; CAA06167.1; -; Genomic DNA.
 DR EMBL; Z72496; CA96577.1; -; Genomic DNA.
 DR EMBL; X74955; CAA52910.1; -; mRNA.
 DR EMBL; U63836; AAB61398.1; -; mRNA.
 DR EMBL; U78554; AAC51344.1; -; Genomic DNA.
 DR EMBL; U78552; AAC51344.1; JOINED; Genomic DNA.
 DR EMBL; U78553; AAC51344.1; JOINED; Genomic DNA.
 DR EMBL; U78551; AAC51343.1; -; mRNA.
 DR EMBL; U95031; AAB65151.1; -; mRNA.
 DR EMBL; Y09788; CAA70926.1; -; Genomic DNA.
 DR Ensembl; ENSG00000117983; Homo sapiens.
 DR HGNC; HGNC:7516; MUC5B.
 DR MIM; 600770; -.
 DR GO; GO:0005576; C:extracellular region; NAS.
 DR GO; GO:0005201; F:extracellular matrix structural constituent; NAS.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR GO; GO:0007155; P:cell adhesion; NAS.
 DR InterPro; IPR006208; Cys_knot.
 DR InterPro; IPR006207; Cys_knot_C.
 DR InterPro; IPR002919; ProtInh_CR_TIL.
 DR InterPro; IPR001007; WVF_C.
 DR InterPro; IPR001846; WVF_D.
 DR Pfam; PF00007; Cys_knot_1.
 DR Pfam; PF01826; TIL; 1.
 DR Pfam; PF00093; WVF; 1.
 DR Pfam; PF00094; WMD; 4.
 DR PROSITE; PS01185; CTCK 1; 1.
 DR PROSITE; PS01225; CTCK 2; 1.
 DR PROSITE; PS01208; WVF_C 1; 2.
 DR PROSITE; PS0184; WVF_C 2; 2.
 DR GlycoProtex; Polymorphin; Repeat; Signal.
 DR SIGNAL 1 25 Potential.
 DR CHAIN 26 5703 Mucin-5B.
 DR DOMAIN 77 225 WFMD 1.
 DR DOMAIN 329 386 TIL.
 DR DOMAIN 426 580 WFMD 2.
 DR DOMAIN 858 918 WFMD 1.
 DR DOMAIN 896 1044 WFMD 3.
 DR DOMAIN 5005 5178 WFMD 4.
 DR DOMAIN 5353 5425 WFMD 2.
 DR DOMAIN 5462 5528 WFMC 3.
 DR DOMAIN 5594 5683 CTCK.
 DR COMPBIAS 1457 1603 Thr-rich.
 DR COMPBIAS 1459 1603 Th-rich.
 DR CARBOHYD 145 145 N-linked (GlcNAc...)
 DR CARBOHYD 201 201 N-linked (GlcNAc...)
 DR CARBOHYD 254 254 N-linked (GlcNAc...)
 DR CARBOHYD 402 402 N-linked (GlcNAc...)
 DR CARBOHYD 516 516 N-linked (GlcNAc...)
 DR CARBOHYD 806 806 N-linked (GlcNAc...)
 DR CARBOHYD 930 930 N-linked (GlcNAc...)
 DR CARBOHYD 1277 1277 N-linked (GlcNAc...)
 DR CARBOHYD 1293 1293 N-linked (GlcNAc...)
 DR CARBOHYD 1557 1557 N-linked (GlcNAc...)
 DR CARBOHYD 1775 1775 N-linked (GlcNAc...)
 DR CARBOHYD 2192 2192 N-linked (GlcNAc...)
 DR CARBOHYD 2721 2721 N-linked (GlcNAc...)
 DR CARBOHYD 3419 3419 N-linked (GlcNAc...)
 DR CARBOHYD 3948 3948 N-linked (GlcNAc...)
 DR CARBOHYD 4745 4745 N-linked (GlcNAc...)
 DR CARBOHYD 4901 4901 N-linked (GlcNAc...)
 DR CARBOHYD 4958 4958 N-linked (GlcNAc...)
 DR CARBOHYD 4965 4965 N-linked (GlcNAc...)
 DR CARBOHYD 4987 4987 N-linked (GlcNAc...)
 DR CARBOHYD 5037 5037 N-linked (GlcNAc...)
 DR CARBOHYD 5052 5052 N-linked (GlcNAc...)
 DR CARBOHYD 5156 5156 N-linked (GlcNAc...)
 DR CARBOHYD 5427 5427 N-linked (GlcNAc...)
 DR CARBOHYD 5467 5467 N-linked (GlcNAc...)
 DR CARBOHYD 5506 5506 N-linked (GlcNAc...)
 DR CARBOHYD 5507 5507 N-linked (GlcNAc...)
 DR CARBOHYD 5543 5543 N-linked (GlcNAc...)
 DR CARBOHYD 5553 5553 N-linked (GlcNAc...)

FT CARBOHYD 5604 5604 N-linked (G1CNAC. . .) (Potential).
 FT CARBOHYD 5618 5618 N-linked (G1CNAC. . .) (Potential).
 FT CARBOHYD 5662 5662 N-linked (G1CNAC. . .) (Potential).
 FT DISULFID 5594 5594 By similarity.
 FT DISULFID 5622 5622 By similarity.
 FT DISULFID 5626 5626 By similarity.
 FT DISULFID 5643 5643 By similarity.
 FT DISULFID 5645 5645 By similarity.
 FT VARIANT 5137 5137 T -> S (in dbSNP:2672788).
 FT CONFLICT 34 34 /FTID=VAR_014123.
 FT CONFLICT 95 100 G -> E (in Ref. 2).
 FT CONFLICT 104 104 FPGCN -> LPCLK (in Ref. 2).
 FT CONFLICT 142 142 S -> C (in Ref. 2).
 FT CONFLICT 225 225 E -> K (in Ref. 1).
 FT CONFLICT 330 330 R -> S (in Ref. 2).
 FT CONFLICT 337 337 PL -> T (in Ref. 2).
 FT CONFLICT 356 356 E -> N (in Ref. 2).
 FT CONFLICT 362 362 E -> K (in Ref. 2).
 FT CONFLICT 369 362 G -> R (in Ref. 2).
 FT CONFLICT 374 374 Missing (in Ref. 2 and 3).
 FT CONFLICT 393 374 D -> N (in Ref. 2).
 FT CONFLICT 468 469 RT -> TR (in Ref. 2).
 FT CONFLICT 512 512 RK -> GR (in Ref. 2).
 FT CONFLICT 585 587 L -> P (in Ref. 2).
 FT CONFLICT 601 601 GAA -> AH (in Ref. 3).
 FT CONFLICT 628 629 A -> S (in Ref. 3).
 FT CONFLICT 633 633 DP -> RS (in Ref. 2).
 FT CONFLICT 676 633 F -> L (in Ref. 2).
 FT CONFLICT 701 676 A -> P (in Ref. 3).
 FT CONFLICT 752 701 R -> P (in Ref. 3).
 FT CONFLICT 752 752 E -> K (in Ref. 2).

Query Match 50.0%; Score 17; DB 1; Length 5703;
 Best Local Similarity 16.7%; Pred. No. 7.1e-09;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
 Db 5062 CTASATMAAARC 5073

RESULT 21
 ID Q6EUA6_ORYSA PRELIMINARY; PRT; 93 AA.
 AC Q6EUA6;
 DT 25-OCT-2004 (TREMUREL. 28, Created)
 DT 25-OCT-2004 (TREMUREL. 28, Last sequence update)
 DE Hypothetical protein QJ1116_C12.20.
 GN Name=QJ1116_C12.20;
 OS Oryza sativa (japonica cultivar-group).
 OC Burkholderia viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Erihartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsunoto T., Yamamoto K.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP004134; BA027763.1; -; Genomic_DNA.
 DR Gramene; Q6EUA6; -;
 KW Hypothetical protein.
 SQ SEQUENCE 93 AA; 9912 MW; 10814C0A65B0F123 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 93;
 Best Local Similarity 16.7%; Pred. No. 1.6e-07;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
 Db 14 CTASAGDATSC 25

RESULT 22
 ID Q62UJ1_ORYSA PRELIMINARY; PRT; 100 AA.
 AC Q62UJ1;
 DT 01-JUN-2003 (TREMUREL. 24, Created)
 DT 01-JUN-2003 (TREMUREL. 24, Last sequence update)
 DT 01-OCT-2003 (TREMUREL. 25, Last annotation update)
 DE Putative topoisomerase.
 GN Name=OSUNBA0032821.11;
 OS Oryza sativa (japonica cultivar-group).
 OC Burkholderia viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Erihartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Eastman A.P., Smith S.C., Bertin N., Liang C., Najjar F.Z., Pratt L.H.,
 RA Cordonnier-Pratt M.-M.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF377947; AA032310.1; -; Genomic_DNA.
 DR Gramene; Q62UJ1; -;
 DR GO; GO:0016853; F:isomerase activity; IBA.
 KW Isomerase.
 SQ SEQUENCE 100 AA; 11026 MW; FA5BF1C9B0965999 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 100;
 Best Local Similarity 16.7%; Pred. No. 1.6e-07;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
 Db 60 CSASAAAAAPC 71

RESULT 23
 ID Q62UJ1_BURMA PRELIMINARY; PRT; 100 AA.
 AC Q62UJ1;
 DT 25-OCT-2004 (TREMUREL. 28, Created)
 DT 25-OCT-2004 (TREMUREL. 28, Last sequence update)
 DT 25-OCT-2004 (TREMUREL. 28, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=BMA1578;
 OS Burkholderia mallei (Pseudomonas mallei).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia.
 OX NCBI_TaxID=13373;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA STRAIN=ATCC 23344;
 RC PubMed=15377793; DOI=10.1073/pnas.040306101;
 RA Nieman W.C., Deshaizer D., Kim H.S., Tettelin H., Nelson K.E.,
 RA Feldblyum T.V., Ulrich R.L., Ronning C.M., Brinkac L.M.,
 RA Daugherty S.C., Davidson T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
 RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
 RA Madupu R., Mahmoud Y., Nelson W.C., Radune D., Romero C.M.,
 RA Sarría S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,
 RA Zafar N., Zhou L., Fraser C.M.;
 RL "Structural flexibility in the Burkholderia mallei genome."
 RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
 DR EMBL; CP000010; AA047930.1; -; Genomic_DNA.
 DR TIGR; BMA1578; -;
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 100 AA; 11000 MW; 89664E4C4C42D9F5 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 100;
 Best Local Similarity 16.7%; Pred. No. 1.6e-07;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
 Db 20 CLAAIAATATATC 31

```

RESULT 24
CBP1_AJECA
ID CBP1_AJECA STANDARD; PRT: 110 AA.
AC 042720;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Calcium-binding protein precursor (CBP).
GN Name=CBP1;
OS Ajellomyces capsulata (Histoplasma capsulatum).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Oxygenales; Ajellomycesaceae; Ajellomyces.
OX NCBI_Taxid=5037;
[1]
RN NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.
RP STRAIN=G186AR;
RC MEDLINE=98149309; PubMed=9489665;
RX Batanghari J.W., Deepe G.S., Jr., Di Cera E., Goldman W.E.;
RA Batanghari J.W., Goldman W.E.;
RT "Calcium dependence and binding in cultures of Histoplasma
RT capsulatum."
RL Infect. Immun. 65:5257-5261(1997).
CC -1- FUNCTION: Involved in calcium binding and uptake in yeast phase.
CC -1- May be important for yeast survival during infection.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- DEVELOPMENTAL STAGE: Expressed in the yeast phase during
CC infection, and not in mycelia.
CC -1- PTM: May be glycosylated.
CC -1- MASS SPECTROMETRY: MW=7858.0; METHOD=Electrospray; RANGE=33-110;
CC NOTE=Ref. 2.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL: AF006209; AAC9354.1; -; Genomic DNA.
KM Calcium; Direct protein sequencing; Glycoprotein; Signal.
FT SIGNAL 1 19 Potential.
FT CHAIN 20 110 Calcium-binding protein.
FT DISULFID 60 93 Potential.
FT DISULFID 65 76 Potential.
FT DISULFID 105 108 Potential.
SQ SEQUENCE 110 AA; 11014 MW; BD98F2FA1B4C77CE CRC64;

Query Match 47.1%; Score 16; DB 1; Length 110;
Best Local Similarity 16.7%; Pred. No. 1.7e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 65 CUSLSASSAAC 76

RESULT 25
Q5C209_SCHA PRELIMINARY; PRT: 144 AA.
ID Q5C209_SCHA PRELIMINARY;
AC Q5C209;
DT 10-MAY-2005 (TReMBLrel. 30, Created)
DT 10-MAY-2005 (TReMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TReMBLrel. 30, Last annotation update)
DE Hypothetical protein.
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;

```

```

OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_Taxid=6182;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Han Z.;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY810427; AAX26316.1; -; mRNA.
DR InterPro: IPR009081; ACP-like.
DR InterPro: IPR011989; ARM-like.
DR InterPro: IPR011992; EF-Hand_type.
KM Hypothetical protein.
SQ SEQUENCE 144 AA; 16762 MW; FF652713F976647B CRC64;

Query Match 47.1%; Score 16; DB 2; Length 144;
Best Local Similarity 16.7%; Pred. No. 1.8e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 15 CSSSYAVSSTSC 26

RESULT 26
Q9YBV0_AERPE PRELIMINARY; PRT: 152 AA.
ID Q9YBV0_AERPE PRELIMINARY;
AC Q9YBV0;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein APE1499.
GN OrderedLocustNames=APE1499;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcates;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_Taxid=56636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KI;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
RA Jin no K., Takahashi M., Sekine M., Baba S.-I., Ankei A., Kosugi H.,
RA Hotoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobid hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix KI."
RL DNA Res. 6:83-101(1999).
DR EMBL: BA000002; BAA80498.1; -; Genomic DNA.
DR PIR: D72630; D72630.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 152 AA; 16241 MW; 4386978A0E6151EF CRC64;

Query Match 47.1%; Score 16; DB 2; Length 152;
Best Local Similarity 16.7%; Pred. No. 1.8e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 137 CSSTATTLLSTRC 148

RESULT 27
Q7SLM9_ORYSA PRELIMINARY; PRT: 161 AA.
ID Q7SLM9_ORYSA PRELIMINARY;
AC Q7SLM9;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Putative DNA topoisomerase (with alternative splicing).
GN Name=OSJNBa0047524.8;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

OC  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC  Euphorbiaceae; Oryzaeae; Oryza.
OX  NCBI_TaxID=39947;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RA  Buehl C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA  Overton II L.L., Tseltrin T., Kim M.M., Bera J.J., Jin S.S.,
RA  Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA  Vanaken S.S., Riedmuller S.B., Utecherack T.T., Feildilyum T.V.,
RA  Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA  White O., Salzberg S.L., Fraser C.M.;
RL  Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  NUCLEOTIDE SEQUENCE.
RA  Buehl R.;
RL  Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AC092556; AAR87258.1; -; Genomic_DNA.
DR  Gramene; Q75LM9; -.
DR  GO; GO:0005524; F:ATP binding; IEA.
DR  GO; GO:0003677; F:DNA binding; IEA.
DR  GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR  GO; GO:0016853; F:isomerase activity; IEA.
DR  GO; GO:0006285; P:DNA topological change; IEA.
DR  InterPro; IPR011991; wing_hlx_DNA_bd.
DR  Pfam; PF04406; TP6A_N; 1.
DR  PRINTS; PR01550; TO6AFAMILY.
KM  DNA-binding; Isomerase; Transcription.
SQ  SEQUENCE 161 AA; 18051 MW; 41B12D72B71B78C3 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 161;
Best Local Similarity 16.7%; Pred. NO. 1.9e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY  4 CXXXXXXXXXXC 15
DB  60 CSSSSSSAGSEC 71

RESULT 28
Q7PY44_ANOGA PRELIMINARY; PRT; 172 AA.
ID  Q7PY44;
AC  Q7PY44;
DT  01-MAR-2004 (TREMBlrel. 26, Created)
DT  01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DE  ENSANGP0000012250 (Fragment).
GN  ORFNames=ENSANG0000009761;
OS  Anopheles gambiae str. PEST.
OC  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preyngota;
OC  Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC  Anophelinae; Anopheles.
OX  NCBI_TaxID=180454;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RA  STRAIN=PEST;
RL  Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC  -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC  -1- CAUTION: The sequence shown here is derived from an
CC  EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC  preliminary data.
DR  EMBL; AAA801069887; EA01398.1; -; Genomic_DNA.
DR  InterPro; IPR001092; HLH_Basic.
DR  Pfam; PF00010; HLH; 1.
DR  PROSITE; PS50888; HLH; 1.
FT  NON TER 1
SQ  SEQUENCE 172 AA; 18464 MW; B857469A135A1884 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 172;
Best Local Similarity 16.7%; Pred. NO. 1.9e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

```

QY  4 CXXXXXXXXXXC 15
DB  77 CSSSSSSAGSEC 88

RESULT 29
Q6ZD90_ORYSA PRELIMINARY; PRT; 176 AA.
ID  Q6ZD90_ORYSA;
AC  Q6ZD90;
DT  05-JUL-2004 (TREMBlrel. 27, Created)
DT  05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT  05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE  Hypothetical protein P0438H08.27.
GN  Name=P0438H08.27;
OS  Oryza sativa (japonica cultivar-group).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC  Euphorbiaceae; Oryzaeae; Oryza.
OX  NCBI_TaxID=39947;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RA  Sasaki T., Matsumoto T., Yamamoto K.;
RL  Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC  -1- SIMILARITY: Contains 1 RING-type zinc finger.
DR  EMBL; AP004460; BAC99511.1; -; Genomic_DNA.
DR  Gramene; Q6ZD90; -.
DR  GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR  GO; GO:0046872; F:metal ion binding; IEA.
DR  GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR  GO; GO:0008270; F:zinc ion binding; IEA.
DR  GO; GO:0016567; P:protein ubiquitination; IEA.
DR  InterPro; IPR01841; Znf_ring.
DR  Pfam; PF00097; zf-C3HC4; 1.
DR  SMART; SM00184; RING; 1.
DR  PROSITE; PS50089; ZF_RING_2; 1.
KM  Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ  SEQUENCE 176 AA; 18608 MW; 79FE9D42C740112 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 176;
Best Local Similarity 16.7%; Pred. NO. 1.9e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY  4 CXXXXXXXXXXC 15
DB  96 CRRRAAAAASAC 107

RESULT 30
Q55014_DICDI PRELIMINARY; PRT; 187 AA.
ID  Q55014_DICDI;
AC  Q55014;
DT  13-SEP-2005 (TREMBlrel. 31, Created)
DT  13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT  13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE  Hypothetical protein.
GN  ORFNames=DD80217903;
OS  Dictyostelium discoideum (Slime mold).
OC  Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX  NCBI_TaxID=44689;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RA  SPRAIN=AX4;
RC  Eichinger L., Pachbat J.A., Gloeckner G., Rajandream M.-A.,
RA  Sugang R., Berriman M., Song J., Olsen R., Szafranski K., Xu O.,
RA  Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA  Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA  Pletcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA  Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bacon N.,
RA  Farbrother P., Desany B., Just B., Morio T., Rost R., Churcher C.,
RA  Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA  Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,
RA  Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA  Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,

```

RA Loulsegged H., Mungall K., Oliver K., Price C., Quail M.A.,
 RA Urushihara H., Hernandez J., Rabinowitch E., Steffen D., Sanders M.,
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
 RA Sngano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
 RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
 RA Shaulsky R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
 RA Williams J., Dear P.H., Noegel A.A., Barrett B., Kuspa A.,
 RT "The genome of the social amoeba *Dictyostelium discoideum*,"
 RL Nature 0:0-0(2005).
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAFI01000042; EAL69047.1; -; Genomic_DNA.
 KM Hypothetical protein.
 SQ SEQUENCE 187 AA; 20124 MW; 8D78DE804DFBC104 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 187;
 Best Local Similarity 16.7%; Pred. No. 1.9e-07;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15
 DB 65 CSYSSSSSSSFFSC 76

RESULT 31

082LKO_STRAW PRELIMINARY; PRT; 208 AA.
 AC 082LKO; 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=SAV2010;
 OS Streptomyces avermectilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxId=33903;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.,
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism *Streptomyces avermectilis*,"
 RL Nat. Biotechnol. 21:526-531(2003).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
 RT "Genome sequence of an industrial microorganism *Streptomyces*
 RT *avermectilis*: deducing the ability of producing secondary
 RT metabolites,"
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL; BA000030; BAC69721.1; -; Genomic_DNA.
 KM Complete proteome; Hypothetical protein.
 SQ SEQUENCE 208 AA; 20876 MW; 4CDCSE8C111AF502 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 208;
 Best Local Similarity 16.7%; Pred. No. 2e-07;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15
 DB 54 CPAASASAAAC 65

RESULT 32
 Q67VV9_ORYSA

ID Q67VV9 ORYSA PRELIMINARY; PRT; 210 AA.
 AC Q67VV9;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
 DE Hypothetical protein OSUNBa0023122.2 (Hypothetical protein
 DE P0530H05.17).
 GN Name=OSUNBa0023122.2; Synonyms=P0530H09.17;
 OS Oryza sativa (japonica cultivar-group),
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzae; Oryza.
 OX NCBI_TaxId=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC
 RT clone:OSUNBa0023122,"
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
 RT clone:P0530H05,"
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP004733; BAD37710.1; -; Genomic_DNA.
 DR EMBL; AP003541; BAD37385.1; -; Genomic_DNA.
 DR Gramene; O67VV9; -;
 KM Hypothetical protein.
 SQ SEQUENCE 210 AA; 21575 MW; C5EB3FF3B9F4457 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 210;
 Best Local Similarity 16.7%; Pred. No. 2e-07;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15
 DB 57 CSTRSSSSSRRC 68

RESULT 33
 Q5K9J0_CRYNE PRELIMINARY; PRT; 214 AA.
 AC Q5K9J0;
 DT 10-MAY-2005 (TREMBlrel. 30, Created)
 DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
 DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=CNK01770;
 OS Cryptococcus neoformans var. neoformans JEC21.
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
 OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
 OX NCBI_TaxId=214684;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=JEC21;
 RA Loftus B., Amedeo P., Roncaglia P., Vamathevan J., Uterback T.,
 RA Van Aken S., Fraser C.;
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=JEC21;
 RX PubMed=15653466; DOI=10.1126/science.1103773;
 RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
 RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
 RA Boesdel I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
 RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
 RA Huang J.C., Jambon G., Jones S.J.M., Koo H.L., Krzyzanski M.I.,
 RA Kwon-Chung K.J., Langelier K.B., Maiti R., Marra M.A., Marra R.E.,
 RA Mathewson C.A., Mitchell T.G., Perlea M., Riggs F.R., Salzberg S.L.,
 RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
 RA Sun B.B., Tenney A., Uterback T.R., Wicks B.L., Wortman J.R.,
 RA Wye N.H., Kronstad J.W., Lodge J.K., Heltman J., Davis R.W.,


```

RA Fraser C.M., Hyman R.W.;
RT "The genome of the basidiomycetous yeast and human pathogen
RT Cryptococcus neoformans.";
RL Science 307:1321-1324(2005).
DR EMBL; AE017351; AAW46337.1; -; Genomic_DNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005732; C:small nuclear ribonucleoprotein complex; IEA.
DR GO; GO:0006397; P:mRNA processing; IEA.
DR InterPro; IPR001163; LSM_snrnp_core.
DR Pfam; PF01423; LSM; 1.
DR SMART; SM00651; Sm; 1.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 214 AA; 23828 MW; 9DEFCE4844BCE42 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 214;
Best Local Similarity 16.7%; Pred. No. 2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 9 CYASTSSVASC 20

RESULT 34
Q55K09 CRYNE PRELIMINARY; PRT; 215 AA.
ID 055K09 CRYNE PRELIMINARY; PRT; 215 AA.
AC 055K09;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DE Hypothetical protein.
GN ORFNames=CNBK1770;
OS Cryptococcus neoformans var. neoformans B-3501A.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=283643;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B-3501A;
RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
RA Wickes B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC Preliminary data.
CC EMBL; AAEY0100052; EAL18156.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 215 AA; 24014 MW; 1CEDFCE4844BCE CRC64;

Query Match 47.1%; Score 16; DB 2; Length 215;
Best Local Similarity 16.7%; Pred. No. 2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 9 CYASTSSVASC 20

RESULT 35
Q5K919 CRYNE PRELIMINARY; PRT; 215 AA.
ID 05K919 CRYNE PRELIMINARY; PRT; 215 AA.
AC 05K919;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DE Hypothetical protein.
GN ORFNames=CNK01770;
OS Cryptococcus neoformans var. neoformans JEC21.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=214684;
RN [1]

```

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=JEC21;
RA Loftus B., Amedeo P., Roncaglia P., Vamathavan J., Uterback T.,
RA Van Aken S., Fraser C.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=JEC21;
RX PubMed=15653466; DOI=10.1126/science.1103773;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathavan J., Miranda M., Anderson I.V., Fraser J.A., Allen J.E.,
RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Jambon G., Jones S.J.M., Koo H.L., Krzywinski M.I.,
RA Kwon-Chung K.J., Lengeler K.B., Maiti R., Marra M.A., Marra R.E.,
RA Matheson C.A., Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L.,
RA Schein J.E., Shwartsbeyn A., Shin H., Shumway M., Specht C.A.,
RA Sub B.B., Tenney A., Uterback T.R., Wickes B.L., Wortman J.R.,
RA Wye N.H., Kronstad J.W., Lodge J.K., Hettman J., Davis R.W.,
RA Fraser C.M., Hyman R.W.;
RT "The genome of the basidiomycetous yeast and human pathogen
RT Cryptococcus neoformans.";
RL Science 307:1321-1324(2005).
DR EMBL; AE017351; AAW46338.1; -; Genomic_DNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005732; C:small nuclear ribonucleoprotein complex; IEA.
DR GO; GO:0006397; P:mRNA processing; IEA.
DR InterPro; IPR001163; LSM_snrnp_core.
DR Pfam; PF01423; LSM; 1.
DR SMART; SM00651; Sm; 1.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 215 AA; 24014 MW; 1CEDFCE4844BCE CRC64;

Query Match 47.1%; Score 16; DB 2; Length 215;
Best Local Similarity 16.7%; Pred. No. 2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 9 CYASTSSVASC 20

RESULT 36
Q4QSJ9 9LECA PRELIMINARY; PRT; 227 AA.
ID 04QSJ9 9LECA PRELIMINARY; PRT; 227 AA.
AC 04QSJ9;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DE Putative non-reducing polyketide synthase (Fragment).
OS Pertusaria coronata.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Lecanoromycetes;
OC Pertusariales; Pertusariaceae; Pertusaria.
OX NCBI_TaxID=232756;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15927215; DOI=10.1016/j.phytochem.2005.04.014;
RA Schmitt I., Martin M.P., Kautz S., Lumbsch H.T.;
RT "Diversity of non-reducing polyketide synthase genes in the
RT Pertusariales (lichenized Ascomycota): A phylogenetic perspective.";
RL Phytochemistry 66:1241-1253(2005).
DR EMBL; AY918748; AAY00086.1; -; Genomic_DNA.
DR InterPro; IPR000794; Ketosacyl_synth.
DR Pfam; PF00109; ketosacyl_synth; 1.
DR PROSITE; PS00606; B_KETOACTL_SYNTHASE; 1.
KW transferase.
FT NON TER 1
FT NON TER 227
SQ SEQUENCE 227 AA; 24445 MW; 4D57BB346B626A8C CRC64;

Query Match 47.1%; Score 16; DB 2; Length 227;
Best Local Similarity 16.7%; Pred. No. 2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

QY 4 CXXXXXXXXXXC 15
DB 95 CSSSPAISTAC 106

RESULT 37
Q7JNB5_GIALA PRELIMINARY; PRT; 230 AA.
ID Q7JNB5;
AC Q7JNB5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Variant-specific surface protein type 4 tsap1/tsa417-like
(Fragment).
GN Name=vsp417-4;
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiae; Giardia.
OX NCBI_TaxId=5741;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9229863; PubMed=10215024; DOI=10.1016/S0166-6851(98)00183-2;
RA Ey P.L., Darby J.M., Mayrhofer G.;
RT "A new locus (vsp417-4) belonging to the tsa417-like subfamily of
RT variant-specific surface protein genes in Giardia intestinalis.";
RL Mol. Biochem. Parasitol. 99:55-68(1999).
DR EMBL: U89153; AAD0041.1; -; Genomic_DNA.
DR InterPro: IPR006212; Furin_repeat.
DR InterPro: IPR005127; Giardia_VSP.
DR Pfam: PF03302; VSP; 1.
DR SMART: SM00181; EGF; 2.
DR SMART: SM00261; FU; 1.
FT NON_TER 1
FT NON_TER 230
SQ SEQUENCE 230 AA; 23336 MW; 45E22ECD1C00B4F1 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 230;
Best Local Similarity 16.7%; Pred. No. 2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 22 CSSTTVRAAATC 33

RESULT 38
Q8LD53_ARATH PRELIMINARY; PRT; 234 AA.
ID Q8LD53_ARATH PRELIMINARY;
AC Q8LD53;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BHLH transcription factor, putative.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22088475; PubMed=12093376;
RA Haas B.U., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN [2]

RP NUCLEOTIDE SEQUENCE.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RA -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.

DR EMBL: AY086197; AAM64276.1; -; mRNA.
DR HSSP: P36956; 1AM9.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0030528; F:transcription regulator activity; IEA.
DR GO: GO:0045449; P:regulation of transcription; IEA.
DR InterPro: IPR001092; HLH_basic.
DR Pfam: PF00010; HLH; 1.
DR SMART: SM00353; HLH; 1.
DR PROSITE: PSS0868; HLH; 1.
SQ SEQUENCE 234 AA; 25470 MW; 9EBEA7D2E44F21D5 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 234;
Best Local Similarity 16.7%; Pred. No. 2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 67 CESSSATSSXAC 78

RESULT 39
Q9FH37_ARATH PRELIMINARY; PRT; 234 AA.
ID Q9FH37_ARATH PRELIMINARY;
AC Q9FH37;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:K5F14
DE (Hypothetical protein At5g54680, K5F14.2) (Putative bHLH transcription
DE factor) (Hypothetical protein At5g54680).
GN Name=At5g54680/K5F14.2; Synonym=At5g54680;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
RN [2]

RP NUCLEOTIDE SEQUENCE.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamuya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [3]

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22598051; PubMed=12679534; DOI=10.1093/molbev/msg088;
RA Heim M.A., Jakob M., Werber M., Martin C., Weisshaar B., Bailey P.C.;
RT "The basic helix-loop-helix transcription factor family in plants: a
RT genome-wide study of protein structure and functional diversity.";
RL Mol. Biol. Evol. 20:735-747(2003).
RN [4]

RP NUCLEOTIDE SEQUENCE.
RA Jakob M.J., Heim M.A., Bailey P., Martin C., Weisshaar B.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [5]

RP NUCLEOTIDE SEQUENCE.
RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamuya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;

RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -1- STABILITY: Contains 1 basic helix-loop-helix (bHLH) domain.
DR EMBL: AF054585; AAK96776.1; -; mRNA.
DR EMBL: AF488629; AAM10964.1; -; mRNA.
DR EMBL: BT002189; AAM2200.1; -; mRNA.
DR HSSP: P36956; 1AM9.
DR GO: GO:0045449; P:regulation of transcription; TAS.
DR InterPro: IPR001092; HLH_basic.
DR Pfam: PF00010; HLH; 1.
DR SMART: SM00353; HLH; 1.
DR PROSITE: PS50888; HLH; 1.
KM Hypothetical protein.
SQ SEQUENCE 234 AA; 25488 MW; 987EA11424F21D1 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 234;
Best Local Similarity 16.7%; Pred. No. 2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15
DB 67 CESSATSSKXC 78

RESULT 40
ID 016421_CAEEL PRELIMINARY; PRT; 244 AA.
AC 016421;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein T05B4.12.
GN ORFNames=T05B4.12;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology";
RL Science 282:2012-2018(1998).
DR EMBL: AF016445; AAC69052.1; -; Genomic_DNA.
DR PIR: T31838; T31838.
DR Ensemble: T05B4.12; Caenorhabditis elegans.
DR WormBase: WBGene00020243; T05B4.12.
DR WormPep: T05B4.12; CE13215.
DR InterPro: IPR003582; SHKT.
DR Pfam: PF01549; SHTK; 3.
DR SMART: SM00254; SHKT; 4.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 244 AA; 25347 MW; 36362492AC861C9 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 244;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15
DB 203 CTSTADSSSTSC 214

RESULT 41
ID 0966B6_CAEEL PRELIMINARY; PRT; 247 AA.
AC 0966B6;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hypothetical protein.

GN ORFNames=Y46H3D.8;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology";
RL Science 282:2012-2018(1998).
DR EMBL: AC006777; AAK72311.1; -; Genomic_DNA.
DR Ensemble: Y46H3D.8; Caenorhabditis elegans.
DR InterPro: IPR003582; SHKT.
DR Pfam: PF01549; SHTK; 3.
DR SMART: SM00254; SHKT; 4.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 247 AA; 25215 MW; FF3CD3BFDD0DE5A7 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 247;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15
DB 206 CTSTADSSSAC 217

RESULT 42
ID 016424_CAEEL PRELIMINARY; PRT; 248 AA.
AC 016424;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein T05B4.3.
GN ORFNames=T05B4.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology";
RL Science 282:2012-2018(1998).
DR EMBL: AF016445; AAC69054.1; -; Genomic_DNA.
DR PIR: T31841; T31841.
DR Ensemble: T05B4.3; Caenorhabditis elegans.
DR WormBase: WBGene00020237; T05B4.3.
DR WormPep: T05B4.3; CE13197.
DR InterPro: IPR003582; SHKT.
DR Pfam: PF01549; SHTK; 3.
DR SMART: SM00254; SHKT; 4.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 248 AA; 25659 MW; 79B281AFB7ED3AFC CRC64;

Query Match 47.1%; Score 16; DB 2; Length 248;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15
DB 207 CTSTADSSSSC 218

RESULT 43
ID 09XV90_CAEEL

ID O9XV90 CAEEL PRELIMINARY; PRT; 257 AA.
AC O9XV90;
DT 01-NOV-1999 (TREMBlrel. 12. Created)
DT 01-NOV-1999 (TREMBlrel. 12. Last sequence update)
DE Hypothetical protein F16H6.3.
GN Name=F16H6.3; ORNames=F16H6.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Bristol N2;
RA Matthews L.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=9906613; PubMed=9851916; DOI=10.1126/science.282.5396.2012;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology."
RL EMBL; Z81506; CAB04129.1; -; Genomic_DNA.
DR PIR; T21029; T21029.
DR Ensemble; F16H6.3; Caenorhabditis elegans.
DR WormBase; WBGene00008893; F16H6.3.
DR WormPep; F16H6.3; C818599.
DR InterPro; IPR003582; SHKT.
DR Pfam; PF01549; SHTK. 3.
DR SMART; SMO0254; SHKT. 4.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 257 AA; 27646 MW; CA204F1AC00EB87 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 257;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXC 15
Db 195 CTNATTTTQASC 206

RESULT 44
O8CAZ5 MOUSE
ID O8CAZ5_MOUSE PRELIMINARY; PRT; 262 AA.
AC O8CAZ5;
DT 01-MAR-2003 (TREMBlrel. 23. Created)
DT 01-MAR-2003 (TREMBlrel. 23. Last sequence update)
DE Mus musculus 6 days neonate skin cDNA, RIKEN full-length enriched
library, clone:A030004108 product:hypothetical protein, full insert
sequence. (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=2108660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai U., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Hara A., Fukunishi Y., Komno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,

RA Saito T., Okazaki Y., Gojobori T., Bond H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochava H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pezole G., Quackenbush J.,
RA Schirai L.M., Scubili F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guetlich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RA The PANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).

RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Komno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama U., Nishi K., Kiteunui T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa K., Izawa M., Ohara E., Watanabe M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multiplexed capillary sequencer."
RL Genome Res. 10:1757-1771(2000).

RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Harada A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hiroane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Komno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohnato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takahashi A., Tanaka T., Tanaka Y., Tanaka Y.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK037168; BAC29730.1; -; mRNA.
DR Ensemble; ENSMUSG0000050239; Mus musculus.
KM Hypothetical protein.
FT NON TER
SQ SEQUENCE 262 AA; 29111 MW; 8D441316CD4432B8 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 262;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 139 CTTATRSASKXC 150

RESULT 45
Q7PRO7 ANOGA PRELIMINARY; PRT; 269 AA.

AC 07PRO7-
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE ENSANGP0000001657 (Fragment).
GN ORFNames=ENSANGC0000001387;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (Apr-2004) to the EMBL/Genbank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA06100847; BAA06779.3; -; Genomic_DNA.
FT NON TER 1 269
FT NON TER 1 269
SQ SEQUENCE 269 AA; 24228 MW; 25BBFF71FD71F1F2 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 269;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 114 CASCSSSTSSC 125

RESULT 46
Q6ZUB7 HUMAN PRELIMINARY; PRT; 278 AA.

AC 06ZUB7-
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein FLJ43840.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakami K., Kanohori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Negahari K., Maeno Y., Nagai K.,
RA Iwagaki T.,
RL Submitted (JUL-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL; AK125828; BAC06309.1; -, mRNA.

SQ SEQUENCE 278 AA; 2768 MW; B0C8757FC1DAE2C0 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 278;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 35 CTAAGSSSLSTC 46

RESULT 47
Q8PF66 XANAC PRELIMINARY; PRT; 285 AA.

AC 08PF66-
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Endonuclease.
GN OrderedLocusNames=XAC4037;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Paron C.S., Furlan L.R.,
RA Queiroz R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camaroto G., Camaroto F., Cardoso J., Chamergo F.,
RA Ciadina L.P., Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorry H., Faria J.B., Ferreira A.J.B., Ferreira R.C.C., Gruber A.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sana J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamara R.E., Teixeira B.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.C.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
DR EMBL; AE012052; AAM3872.1; -; Genomic_DNA.
DR HSSP; P13717; 1G8T.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001604; Endonuclease.
DR Pfam; PF01223; Endonuclease_NS; 1.
DR SMART; SM00477; NUC; 1.
DR PROSITE; PS01070; NUCLEASE_NON_SPEC; UNKNOWN_1.
KW Complete proteome; Endonuclease.
SQ SEQUENCE 285 AA; 30358 MW; 5581A0E9A2057BB CRC64;

Query Match 47.1%; Score 16; DB 2; Length 285;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 51 CTSVAATTRC 62

RESULT 48
Q6ZQ01 HUMAN PRELIMINARY; PRT; 286 AA.

AC 06ZQ01-
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

```

DE Hypothetical protein FLJ46328.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;
OC Homo.
CX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Testis;
RA Nishimura K., Nagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuka T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Sato K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikkawa H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshino A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuko Y., Nagai K., Isegai T.,
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128833; BAC87633.1; -; mRNA.
SQ SEQUENCE 286 AA; 30415 MW; B6CEFAFA76737B8C CRC64;

Query Match 47.1%; Score 16; DB 2; Length 286;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 231 CTAAGSSSLSTC 242

RESULT 49
Q9F934_BIFAD
ID Q9F934_BIFAD PRELIMINARY; PRT; 287 AA.
AC Q9F934;
DT 01-MAR-2001 (Tremblrel. 16, Last Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Sugar binding protein Sbp.
GN Name=sbp;
OS Bifidobacterium adolescentis.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
CX NCBI_TaxID=1680;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=INT57;
RA Park M.S., Yoon H.J., Ji G.E.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF213175; AAG31696.1; -; Genomic DNA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR006059; SBP_bac_1.
DR Pfam; PF01547; SBP_bac_1; 1.
SQ SEQUENCE 287 AA; 32054 MW; 11B9382AE39B86DB CRC64;

Query Match 47.1%; Score 16; DB 2; Length 287;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 48 CAAIMAASAAAC 59

RESULT 50
Q89DY5_BRAJA
ID Q89DY5_BRAJA PRELIMINARY; PRT; 297 AA.
AC Q89DY5;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE B117302 protein.
GN OrderedLocustNames=b117302;

```

```

OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
CX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=2248498; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamide K., Uehimi T.,
RA Sasanaka S., Watanabe A., Idegawa K., Itiguchi M., Kawashima K.,
RA Kohara M., Matsunoto M., Shimo S., Tsunoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000040; BACS2567.1; -; Genomic DNA.
KW Complete proteome.
SQ SEQUENCE 297 AA; 32798 MW; E8E4D21866645601 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 297;
Best Local Similarity 16.7%; Pred. No. 2.2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 250 CSATHSSDAAC 261

```

Search completed: January 4, 2006, 16:09:50
 Job time : 143.548 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 : Search time 54.7478 Seconds
(without alignments)
64,204 Million cell updates/sec

Title: US-09-932-322-9
Perfect score: 24
Sequence: 1 CXXXXXXC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A Geneseq_21:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*
- 9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	75.0	9	ADQ13022	Adq13022 Hepatitis
2	18	75.0	9	ADQ13021	Adq13021 Hepatitis
3	18	75.0	11	AAW65943	AAW65943 Molecule
4	18	75.0	11	AAW65944	AAW65944 Molecule
5	18	75.0	15	ADV23688	Adv23688 HBV Immun
6	18	75.0	15	ADV23687	Adv23687 HBV Immun
7	18	75.0	24	ABR42273	Abbr42273 Peptide #
8	18	75.0	24	AAW36081	AAW36081 Peptide #
9	18	75.0	24	AAW75973	AAW75973 Human bon
10	18	75.0	24	AAW63159	AAW63159 Human bon
11	18	75.0	24	ABG57702	Abg57702 Human liv
12	18	75.0	69	ABR68695	Abbr68695 Drosophil
13	18	75.0	69	ADP30715	Adp30715 Human sec
14	18	75.0	91	ABR39058	Abbr39058 Peptide #
15	18	75.0	91	AAW32549	AAW32549 Peptide #
16	18	75.0	91	AAW72290	AAW72290 Human bon
17	18	75.0	91	AAW59710	AAW59710 Human bra
18	18	75.0	91	ABG53976	Abg53976 Human liv
19	18	75.0	91	ABG42105	Abg42105 Human pep
20	18	75.0	93	ADP30859	Adp30859 Human sec
21	18	75.0	99	ABO83167	Abob83167 Pseudomon
22	18	75.0	99	ABO30709	Abob30709 Human sec
23	18	75.0	107	ABO74878	Abob74878 Pseudomon
24	18	75.0	115	ABO73394	Abob73394 Pseudomon

25	18	75.0	122	7	ADP04777	Adt04777 Bacterial
26	18	75.0	135	8	ADP30703	Adp30703 Human sec
27	18	75.0	144	7	ABO75730	Abob75730 Pseudomon
28	18	75.0	144	8	ADP31453	Adp31453 Human sec
29	18	75.0	144	8	ADP90450	Adp90450 Plant ful
30	18	75.0	150	6	ABU45351	Abu45351 Protein e
31	18	75.0	153	3	ABR41043	Abbr41043 Human ORF
32	18	75.0	153	5	ABP09251	Abp09251 Human ORF
33	18	75.0	154	7	ABO74421	Abob74421 Pseudomon
34	18	75.0	156	7	ABO77836	Abob77836 Pseudomon
35	18	75.0	159	8	ADP30761	Adp30761 Human sec
36	18	75.0	162	8	ADP30758	Adp30758 Human sec
37	18	75.0	162	8	ADP30756	Adp30756 Human sec
38	18	75.0	162	8	ADP30757	Adp30757 Human sec
39	18	75.0	162	8	ADP30813	Adp30813 Human sec
40	18	75.0	165	8	ADP31269	Adp31269 Human sec
41	18	75.0	168	8	ADP31099	Adp31099 Human sec
42	18	75.0	171	8	ADP30794	Adp30794 Human sec
43	18	75.0	171	3	ADP30793	Adp30793 Human sec
44	18	75.0	177	3	AAQ10193	Aaq10193 Arabidops
45	18	75.0	177	8	ADP30755	Adp30755 Human sec
46	18	75.0	177	8	ADP31682	Adp31682 Human sec
47	18	75.0	177	8	ADP56379	Adp56379 Plant pol
48	18	75.0	182	7	ABM86593	Abm86593 Rice abio
49	18	75.0	185	7	ABM89559	Abm89559 Rice abio
50	18	75.0	192	8	ADP31576	Adp31576 Human sec
51	18	75.0	194	7	ABO78166	Abob78166 Pseudomon
52	18	75.0	199	8	ADY13615	Ady13615 Plant ful
53	18	75.0	201	8	ADP30711	Adp30711 Human sec
54	18	75.0	204	3	AAQ10192	Aaq10192 Arabidops
55	18	75.0	210	8	ADP30986	Adp30986 Human sec
56	18	75.0	210	8	ADP30960	Adp30960 Human sec
57	18	75.0	219	8	ADP30593	Adp30593 Human sec
58	18	75.0	222	8	ADP31354	Adp31354 Human sec
59	18	75.0	228	8	ADP30764	Adp30764 Human sec
60	18	75.0	233	3	AAV74791	Aav74791 Neisseria
61	18	75.0	234	8	ADP31480	Adp31480 Human sec
62	18	75.0	234	8	ADP31481	Adp31481 Human sec
63	18	75.0	237	8	ADP30792	Adp30792 Human sec
64	18	75.0	239	7	ABO81291	Abob81291 Pseudomon
65	18	75.0	249	8	ADP30754	Adp30754 Human sec
66	18	75.0	249	8	ADP31443	Adp31443 Human sec
67	18	75.0	252	8	ADP31485	Adp31485 Human sec
68	18	75.0	258	7	ABO80689	Abob80689 Pseudomon
69	18	75.0	258	8	ADP30597	Adp30597 Human sec
70	18	75.0	267	8	ADP30822	Adp30822 Human sec
71	18	75.0	270	7	ABO68843	Abob68843 Pseudomon
72	18	75.0	276	8	ADP30568	Adp30568 Human sec
73	18	75.0	277	8	ADP30503	Adp30503 Human sec
74	18	75.0	279	8	ADP31107	Adp31107 Human sec
75	18	75.0	291	8	ADP31193	Adp31193 Human sec
76	18	75.0	294	8	ADP31045	Adp31045 Human sec
77	18	75.0	300	8	ADP30775	Adp30775 Human sec
78	18	75.0	309	8	ADP30862	Adp30862 Human sec
79	18	75.0	309	8	ADP30873	Adp30873 Human sec
80	18	75.0	310	7	ABO80230	Abob80230 Pseudomon
81	18	75.0	312	8	ADP31454	Adp31454 Human sec
82	18	75.0	318	7	ABM85364	Abm85364 Mouse pro
83	18	75.0	318	8	ADP30588	Adp30588 Human sec
84	18	75.0	318	8	ADP30824	Adp30824 Human sec
85	18	75.0	320	8	ADP31607	Adp31607 Human sec
86	18	75.0	320	8	ADP31649	Adp31649 Human sec
87	18	75.0	323	7	ABO74543	Abob74543 Pseudomon
88	18	75.0	327	7	ABM89959	Abm89959 Rice abio
89	18	75.0	330	4	ABR59598	Abbr59598 Drosophil
90	18	75.0	330	8	ADP30892	Adp30892 Human sec
91	18	75.0	339	8	ADP30702	Adp30702 Human sec
92	18	75.0	345	8	ADP31016	Adp31016 Human sec
93	18	75.0	345	8	ADP31683	Adp31683 Human sec
94	18	75.0	348	8	ADP58550	Adp58550 Plant pol
95	18	75.0	348	8	ADP31676	Adp31676 Human sec
96	18	75.0	348	8	ADP31374	Adp31374 Human sec
97	18	75.0	353	8	ADP31558	Adp31558 Human sec

98	18	75.0	354	8	ADP30783	Human sec	171	18	75.0	558	8	ADP31255	ADP31255 Human sec
99	18	75.0	354	8	ADP30707	Human sec	172	18	75.0	564	8	ADP31194	ADP31194 Human sec
100	18	75.0	354	8	ADP30779	Human sec	173	18	75.0	574	8	ADP31133	ADP31133 Human sec
101	18	75.0	354	8	ADP30784	Human sec	174	18	75.0	576	8	ADP31248	ADP31248 Human sec
102	18	75.0	355	4	ADP30778	Human sec	175	18	75.0	583	2	ADH11358	ADH11358 Human sec
103	18	75.0	355	4	ABM57739	Drosophila	176	18	75.0	588	8	ADP31680	ADP31680 Human sec
104	18	75.0	355	7	ABM87551	Rice abio	177	18	75.0	592	8	ADP30917	ADP30917 Human sec
105	18	75.0	355	9	ADZ76101	Fruit fly	178	18	75.0	594	8	ADP30895	ADP30895 Human sec
106	18	75.0	357	8	ADP31223	Human sec	179	18	75.0	598	8	ADH11789	ADH11789 Human the
107	18	75.0	357	8	ADP31267	Human sec	180	18	75.0	603	8	ADP31150	ADP31150 Human sec
108	18	75.0	360	8	ADP31439	Human sec	181	18	75.0	604	8	ADP30940	ADP30940 Human sec
109	18	75.0	366	8	ADP31106	Human sec	182	18	75.0	604	8	ADP30941	ADP30941 Human sec
110	18	75.0	366	8	ADP31670	Human sec	183	18	75.0	605	8	ADP30507	ADP30507 Human sec
111	18	75.0	367	5	AAO21694	Human sec	184	18	75.0	609	8	ADP31263	ADP31263 Human sec
112	18	75.0	369	7	ABM88560	Rice abio	185	18	75.0	612	8	ADP31064	ADP31064 Human sec
113	18	75.0	370	8	ADY09767	Plant ful	186	18	75.0	615	8	ADP31132	ADP31132 Human sec
114	18	75.0	372	8	ADP31610	Human sec	187	18	75.0	615	8	ADP31361	ADP31361 Human sec
115	18	75.0	382	8	ADP31216	Human sec	188	18	75.0	618	8	ADP31360	ADP31360 Human sec
116	18	75.0	382	8	ADP31221	Human sec	189	18	75.0	617	8	ADP31657	ADP31657 Human sec
117	18	75.0	387	8	ADP31158	Human sec	190	18	75.0	618	8	ADP31554	ADP31554 Human sec
118	18	75.0	390	8	ADP31218	Human sec	191	18	75.0	621	8	ADP30896	ADP30896 Human sec
119	18	75.0	393	8	ADP31345	Human sec	192	18	75.0	639	8	ADP31521	ADP31521 Human sec
120	18	75.0	399	8	ADP31464	Human sec	193	18	75.0	639	8	ADP30861	ADP30861 Human sec
121	18	75.0	401	7	ABO68344	Pseudomon	194	18	75.0	642	8	ADP31265	ADP31265 Human sec
122	18	75.0	411	8	ADP31104	Human sec	195	18	75.0	645	8	ADP31124	ADP31124 Human sec
123	18	75.0	411	8	ADP30729	Human sec	196	18	75.0	645	8	ADP31125	ADP31125 Human sec
124	18	75.0	420	8	ADP31349	Human sec	197	18	75.0	645	8	ADP30858	ADP30858 Human sec
125	18	75.0	421	8	ADP31159	Human sec	198	18	75.0	645	8	ADP31183	ADP31183 Human sec
126	18	75.0	423	8	ADP31479	Human sec	199	18	75.0	654	8	ADP30666	ADP30666 Human sec
127	18	75.0	425	8	ADP30549	Human sec	200	18	75.0	658	8	ADP31226	ADP31226 Human sec
128	18	75.0	429	8	ADP31207	Human sec	201	18	75.0	665	8	ADP30571	ADP30571 Human sec
129	18	75.0	434	5	ABP41972	Human cva	202	18	75.0	666	8	ADP31547	ADP31547 Human sec
130	18	75.0	441	8	ADP31206	Human sec	203	18	75.0	666	8	ADP31256	ADP31256 Human sec
131	18	75.0	449	8	ADP31283	Human sof	204	18	75.0	669	8	ADP31598	ADP31598 Human sec
132	18	75.0	450	8	ADP30887	Human sec	205	18	75.0	669	8	ADP31493	ADP31493 Human sec
133	18	75.0	453	8	ADP31465	Human sec	206	18	75.0	677	8	ADP31585	ADP31585 Human sec
134	18	75.0	471	8	ADP31567	Human sec	207	18	75.0	681	8	ADP30868	ADP30868 Human sec
135	18	75.0	478	8	ADP31007	Human sec	208	18	75.0	681	8	ADP31053	ADP31053 Human sec
136	18	75.0	483	8	ADP30852	Human sec	209	18	75.0	699	8	ADP31080	ADP31080 Human sec
137	18	75.0	483	8	ADP30853	Human sec	210	18	75.0	699	8	ADP31079	ADP31079 Human sec
138	18	75.0	492	4	AAH71707	Human erf	211	18	75.0	699	8	ADP31241	ADP31241 Human sec
139	18	75.0	492	7	ADD14170	Human src	212	18	75.0	699	8	ADP31240	ADP31240 Human sec
140	18	75.0	494	6	ABR64215	Angiogene	213	18	75.0	702	8	ADP31518	ADP31518 Human sec
141	18	75.0	494	8	ADO19178	Human PRO	214	18	75.0	711	8	ADP31215	ADP31215 Human sec
142	18	75.0	494	8	ADO19174	Human PRO	215	18	75.0	711	8	ADP31652	ADP31652 Human sec
143	18	75.0	494	8	ADP54221	Human PRO	216	18	75.0	711	8	ADP31535	ADP31535 Human sec
144	18	75.0	494	9	ADX05928	Cyclin-de	217	18	75.0	711	8	ADP31440	ADP31440 Human sec
145	18	75.0	494	9	ADY14502	PRO polyP	218	18	75.0	714	8	ADP31561	ADP31561 Human sec
146	18	75.0	494	9	ADY14498	PRO polyP	219	18	75.0	720	5	AAO14994	AAO14994 Lamtln1-r
147	18	75.0	502	7	ADC35086	Human bre	220	18	75.0	720	8	ADW47628	ADW47628 Lamtln1-5
148	18	75.0	513	8	ADP31681	Human sec	221	18	75.0	722	8	ADP98843	ADP98843 C. albica
149	18	75.0	514	8	ADP31122	Human sec	222	18	75.0	750	8	ADP31131	ADP31131 Human sec
150	18	75.0	516	7	ABM85365	Human pro	223	18	75.0	756	8	ADP31039	ADP31039 Human sec
151	18	75.0	516	8	ADP30871	Human sec	224	18	75.0	757	8	ADP30925	ADP30925 Human sec
152	18	75.0	516	8	ADP31418	Human sec	225	18	75.0	765	8	ADP31149	ADP31149 Human sec
153	18	75.0	531	8	ADP31539	Human sec	226	18	75.0	768	8	ADP30912	ADP30912 Human sec
154	18	75.0	531	8	ADP30594	Human sec	227	18	75.0	768	8	ADP31362	ADP31362 Human sec
155	18	75.0	531	8	ADP31540	Human sec	228	18	75.0	768	8	ADP31363	ADP31363 Human sec
156	18	75.0	531	8	ADP31696	Human sec	229	18	75.0	771	8	ADP30907	ADP30907 Human sec
157	18	75.0	537	8	ADP31557	Human sec	230	18	75.0	771	8	ADP31469	ADP31469 Human sec
158	18	75.0	538	8	ADY22481	Plant ful	231	18	75.0	771	8	ADP31244	ADP31244 Human sec
159	18	75.0	541	2	ADH11359	Vertebrat	232	18	75.0	774	8	ADP30885	ADP30885 Human sec
160	18	75.0	543	8	ADP30864	Human sec	233	18	75.0	774	8	ADP30506	ADP30506 Human sec
161	18	75.0	546	8	ADO19176	Human PRO	234	18	75.0	774	8	ADP31225	ADP31225 Human sec
162	18	75.0	546	8	ADP31449	Human sec	235	18	75.0	779	8	ADP30915	ADP30915 Human sec
163	18	75.0	546	8	ADP54223	Human PRO	236	18	75.0	779	8	ADP30897	ADP30897 Human sec
164	18	75.0	546	9	ADY14500	PRO polyP	237	18	75.0	780	8	ADP31077	ADP31077 Human sec
165	18	75.0	549	8	ADP31009	Human sec	238	18	75.0	783	8	ADP31436	ADP31436 Human sec
166	18	75.0	552	8	ADP30875	Human sec	239	18	75.0	783	8	ADP30496	ADP30496 Human sec
167	18	75.0	552	8	ADP31019	Human sec	240	18	75.0	786	8	ADP30494	ADP30494 Human sec
168	18	75.0	555	8	ADP31168	Human sec	241	18	75.0	795	8	ADP31332	ADP31332 Human sec
169	18	75.0	557	8	ADP31103	Human sec	242	18	75.0	804	8	ADP31635	ADP31635 Human sec
170	18	75.0	558	8	ADP31254	Human sec	243	18	75.0	804	8	ADP31291	ADP31291 Human sec

536	18	75.0	4440	6	ABU95207	Novel	hum	609	17	70.8	53	6	ABM54353	Abm54353	Propionib
537	18	75.0	4440	6	ABU90755	Novel	hum	610	17	70.8	53	6	ABM53320	Abm53320	Propionib
538	18	75.0	4440	6	ABU93917	Novel	hum	611	17	70.8	54	6	AEA30185	Aea30185	Pertussis
539	18	75.0	4440	6	ABU86191	Novel	hum	612	17	70.8	58	4	AAU54758	AAU54758	Propionib
540	18	75.0	4440	6	ABU82046	Novel	hum	613	17	70.8	58	6	ABM51277	Abm51277	Propionib
541	18	75.0	4440	6	ABU07907	Novel	hum	614	17	70.8	58	9	ADZ11761	Adz11761	B. coli D
542	18	75.0	4440	6	ABU94227	Novel	hum	615	17	70.8	58	9	AEA30433	Aea30433	Pertussis
543	18	75.0	4440	6	ABO00100	Novel	hum	616	17	70.8	58	9	AEA30308	Aea30308	Pertussis
544	18	75.0	4440	6	ABU87111	Novel	hum	617	17	70.8	59	4	AAU51154	AAU51154	Propionib
545	18	75.0	4440	6	ABU91352	Novel	hum	618	17	70.8	59	6	ABM47673	Abm47673	Propionib
546	18	75.0	4440	6	ABU90445	Novel	hum	619	17	70.8	60	9	AEA30430	Aea30430	Pertussis
547	18	75.0	4440	6	ABU97036	Novel	hum	620	17	70.8	60	9	AEA30259	Aea30259	Pertussis
548	18	75.0	4440	6	ABO05232	Novel	hum	621	17	70.8	60	9	AEA30493	Aea30493	Pertussis
549	18	75.0	4683	8	ADP31260	Human	sec	622	17	70.8	60	9	AEA30376	Aea30376	Pertussis
550	18	75.0	4752	8	ADP30585	Human	sec	623	17	70.8	60	9	AEA30494	Aea30494	Pertussis
551	18	75.0	4752	8	ADP30651	Human	sec	624	17	70.8	60	9	AEA30431	Aea30431	Pertussis
552	18	75.0	4834	8	ADP31259	Human	sec	625	17	70.8	60	9	AEA30432	Aea30432	Pertussis
553	18	75.0	5304	8	ADP30706	Human	sec	626	17	70.8	62	7	ADU38115	Adu38115	Human par
554	18	75.0	5397	8	ADP31068	Human	sec	627	17	70.8	62	7	ABO69433	AbO69433	Pseudomon
555	18	75.0	5514	8	ADP31186	Human	sec	628	17	70.8	64	2	AAV12328	AAV12328	Human 5'
556	18	75.0	5514	8	ADP31591	Human	sec	629	17	70.8	64	2	AAW67942	AAW67942	Fragment
557	18	75.0	5820	8	ADP31118	Human	sec	630	17	70.8	64	4	AAW85119	AAW85119	Human imm
558	18	75.0	6465	8	ADP30705	Human	sec	631	17	70.8	64	5	ABP69086	Abp69086	Human pol
559	18	75.0	6729	8	ADP31600	Human	sec	632	17	70.8	66	4	ABG13638	ABG13638	Novel hum
560	18	75.0	7285	6	ABJ38280	DAMG21 -RA		633	17	70.8	67	7	ABM86677	ABM86677	Rice
561	18	75.0	8973	6	AAO16358	Human	tra	634	17	70.8	68	7	ABO09377	ABO09377	Human ORF
562	18	75.0	8973	6	ADP31119	Human	sec	635	17	70.8	68	7	ABO78891	ABO78891	Pseudomon
563	18	75.0	8976	8	ADP31425	Human	sec	636	17	70.8	70	4	AAU53345	AAU53345	Propionib
564	18	75.0	9195	8	ADP31494	Human	sec	637	17	70.8	70	6	ABM49864	ABM49864	Propionib
565	18	75.0	10944	8	ADP31311	Human	sec	638	17	70.8	71	2	AAW83934	AAW83934	Human sec
566	18	75.0	11328	8	ADP31310	Human	sec	639	17	70.8	71	3	AAH34715	AAH34715	Human sec
567	17	70.8	8	5	ABB47073	Desmocoll		640	17	70.8	73	7	ABO82698	ABO82698	Pseudomon
568	17	70.8	8	5	ABB46610	Desmocoll		641	17	70.8	73	8	ADP29540	ADP29540	Human sec
569	17	70.8	8	5	ADZ50990	Amino aci		642	17	70.8	74	4	AAU59878	AAU59878	Propionib
570	17	70.8	8	9	ADZ50974	Amino aci		643	17	70.8	74	6	ABM56397	ABM56397	Propionib
571	17	70.8	10	4	AAW95975	Human com		644	17	70.8	77	9	ADV76639	ADV76639	Human TER
572	17	70.8	10	4	AAW95917	Human com		645	17	70.8	78	9	ADW73420	ADW73420	Plant ful
573	17	70.8	20	2	AAW23344	Peptide s		646	17	70.8	78	9	AAW94485	AAW94485	Human
574	17	70.8	20	3	AAE19697	O-Methylt		647	17	70.8	79	4	AAW19029	AAW19029	Peptide #
575	17	70.8	20	4	AAE05831	Liquidamb		648	17	70.8	79	4	ABR38235	ABR38235	Peptide #
576	17	70.8	20	9	ADZ14872	Liquidamb		649	17	70.8	79	4	AAW31668	AAW31668	Peptide #
577	17	70.8	22	4	AAW33651	Peptide #		650	17	70.8	79	4	ABR23413	ABR23413	Protein
578	17	70.8	22	4	AAW73453	Human bon		651	17	70.8	79	4	ABG53088	ABG53088	Human liv
579	17	70.8	22	4	AAW60779	Human bon		652	17	70.8	79	5	ABG41186	ABG41186	Human pep
580	17	70.8	22	4	ABG55176	Human liv		653	17	70.8	80	3	AAW11389	AAW11389	Arbidops
581	17	70.8	22	5	ABG43312	Human pep		654	17	70.8	83	4	AAW64692	AAW64692	Propionib
582	17	70.8	24	8	ADP73753	Influenza		655	17	70.8	83	6	ABM61211	ABM61211	Propionib
583	17	70.8	28	9	AAW98776	Human cel		656	17	70.8	84	3	AAW60503	AAW60503	Arbidops
584	17	70.8	28	9	ADZ47463	Human HAS		657	17	70.8	84	3	AAW61337	AAW61337	Propionib
585	17	70.8	34	5	AAU87022	Adenoviru		658	17	70.8	84	6	AAW57856	AAW57856	Propionib
586	17	70.8	35	6	ABU61319	Human A d		659	17	70.8	85	9	ABM41187	ABM41187	L. pneumo
587	17	70.8	35	6	ABU61408	Low densi		660	17	70.8	85	9	AAU51874	AAU51874	Propionib
588	17	70.8	35	8	ADP21541	Human LDL		661	17	70.8	86	4	AAW48393	AAW48393	Propionib
589	17	70.8	37	4	AAO09823	Human pol		662	17	70.8	86	6	AAW63689	AAW63689	Propionib
590	17	70.8	39	9	ABE33285	Human LRP		663	17	70.8	87	4	ABM60208	ABM60208	Propionib
591	17	70.8	39	9	AAU59877	Human LRP		664	17	70.8	87	6	ADP17451	ADP17451	Mouse IL-
592	17	70.8	50	6	ABM56396	Propionib		665	17	70.8	87	7	AAU21754	AAU21754	Novel hum
593	17	70.8	51	4	ABR39045	Peptide #		666	17	70.8	88	4	AAU45755	AAU45755	Propionib
594	17	70.8	51	4	AAW32534	Peptide #		667	17	70.8	88	4	AAU20119	AAU20119	Human DNA
595	17	70.8	51	4	AAW72275	Human bon		668	17	70.8	88	5	ABG91368	ABG91368	Novel hum
596	17	70.8	51	4	AAW59697	Human bra		669	17	70.8	88	5	ABM42274	ABM42274	Propionib
597	17	70.8	51	4	ABG53961	Human liv		670	17	70.8	88	7	ADC46395	ADC46395	Human neo
598	17	70.8	51	5	ABM42090	Human pep		671	17	70.8	88	7	ABO69228	ABO69228	Pseudomon
599	17	70.8	53	4	ABM19878	Peptide #		672	17	70.8	88	7	ADP31655	ADP31655	Human sec
600	17	70.8	53	4	ABR39811	Peptide #		673	17	70.8	90	8	AAV12296	AAV12296	Human 5'
601	17	70.8	53	4	AAW33401	Protein #		674	17	70.8	92	5	ABO09944	ABO09944	Guinea pi
602	17	70.8	53	4	ABR24424	Protein #		675	17	70.8	92	5	ABM09943	ABM09943	Human int
603	17	70.8	53	4	AAU56801	Propionib		676	17	70.8	92	7	AAE33936	AAE33936	Guinea pi
604	17	70.8	53	4	AAU57834	Propionib		677	17	70.8	92	7	AAE33935	AAE33935	Human int
605	17	70.8	53	4	ABG54907	Human liv		678	17	70.8	92	9	AAE35800	AAE35800	Partial i
606	17	70.8	53	4	ABG43038	Human pep		679	17	70.8	92	9	AAU56711	AAU56711	Propionib
607	17	70.8	53	5				680							
608	17	70.8	53	5				681							

682	17	70.8	93	6	ABM53230	755	17	70.8	134	6	ABM55230	Adm55230	Propionib
683	17	70.8	96	6	AAU52513	756	17	70.8	134	7	ABO76893	ABO76893	Pseudomon
684	17	70.8	96	6	ABM49032	757	17	70.8	135	3	AAV75297	AAV75297	Neisseria
685	17	70.8	97	4	AAU66144	758	17	70.8	135	3	AAV75296	AAV75296	Neisseria
686	17	70.8	97	5	ABP09938	759	17	70.8	135	5	ABP08959	ABP08959	Human ORF
687	17	70.8	97	6	ABM62663	760	17	70.8	135	7	ABO74109	ABO74109	Pseudomon
688	17	70.8	97	7	ABM62645	761	17	70.8	135	8	ADP31638	ADP31638	Human sec
689	17	70.8	97	9	ABO79445	762	17	70.8	135	8	ADY07492	ADY07492	Plant ful
690	17	70.8	99	7	ABE37882	763	17	70.8	136	3	AAV75295	AAV75295	Neisseria
691	17	70.8	99	4	AAU49262	764	17	70.8	136	7	ADM57545	ADM57545	Human his
692	17	70.8	99	5	ABP43926	765	17	70.8	136	7	ABO73988	ABO73988	Pseudomon
693	17	70.8	99	6	ABM45781	766	17	70.8	136	7	ABM66687	ABM66687	Rice abio
694	17	70.8	102	7	ADN04570	767	17	70.8	137	3	AAV65413	AAV65413	Human 5'
695	17	70.8	105	7	ABO79727	768	17	70.8	137	8	ADU72977	ADU72977	Signal pe
696	17	70.8	105	7	ADBE2627	769	17	70.8	137	9	ADZ73968	ADZ73968	Human com
697	17	70.8	106	4	AAAG64065	770	17	70.8	140	7	ADFI3949	ADFI3949	Human end
698	17	70.8	106	4	ABM15333	771	17	70.8	140	7	ABO82130	ABO82130	Pseudomon
699	17	70.8	106	5	AAU96154	772	17	70.8	141	7	ABO77083	ABO77083	Pseudomon
700	17	70.8	106	6	ADA57463	773	17	70.8	143	5	AAU96140	AAU96140	Human BRP
701	17	70.8	106	6	ADA41343	774	17	70.8	144	8	ABO79215	ABO79215	Pseudomon
702	17	70.8	106	6	ABR48065	775	17	70.8	144	8	ADI39724	ADI39724	TADG-14 p
703	17	70.8	106	6	ABP72423	776	17	70.8	144	8	ADI37148	ADI37148	Human pro
704	17	70.8	106	8	ABP72423	777	17	70.8	144	8	ADP31474	ADP31474	Human sec
705	17	70.8	106	9	ADW86161	778	17	70.8	148	7	ADFI3954	ADFI3954	Human end
706	17	70.8	107	7	ADFI13950	779	17	70.8	148	7	ABO74263	ABO74263	Human gen
707	17	70.8	110	8	ADP30723	780	17	70.8	149	4	AAU68926	AAU68926	Human pro
708	17	70.8	111	4	ABG03591	781	17	70.8	149	7	ADE31059	ADE31059	Human dia
709	17	70.8	112	7	ABO77337	782	17	70.8	149	7	ADE13246	ADE13246	Protease
710	17	70.8	113	7	AAU50472	783	17	70.8	149	8	ADH78455	ADH78455	Human pro
711	17	70.8	113	6	ABM46991	784	17	70.8	150	8	ADX66524	ADX66524	Plant ful
712	17	70.8	113	8	ADP31617	785	17	70.8	154	7	ABO74102	ABO74102	Pseudomon
713	17	70.8	113	8	ADP31613	786	17	70.8	154	7	ABO60522	ABO60522	Human gen
714	17	70.8	117	3	AAAB25477	787	17	70.8	156	7	ABO80146	ABO80146	Pseudomon
715	17	70.8	117	7	ADBE5158	788	17	70.8	156	8	ADP30759	ADP30759	Human sec
716	17	70.8	117	7	ABO68458	789	17	70.8	157	8	ADX78826	ADX78826	Plant ful
717	17	70.8	118	6	ABU99121	790	17	70.8	162	3	AAI95932	AAI95932	Porcine a
718	17	70.8	118	8	ADM93807	791	17	70.8	162	4	AAU42696	AAU42696	Pseudomon
719	17	70.8	120	7	ADBE4538	792	17	70.8	162	6	ABM39215	ABM39215	Propionib
720	17	70.8	120	8	ADP30556	793	17	70.8	163	7	ABO80365	ABO80365	Pseudomon
721	17	70.8	122	4	AAU42386	794	17	70.8	163	7	ABO68705	ABO68705	Pseudomon
722	17	70.8	122	6	ABM38905	795	17	70.8	164	7	ABO78378	ABO78378	Pseudomon
723	17	70.8	123	7	ABO66168	796	17	70.8	165	4	AAU58075	AAU58075	Propionib
724	17	70.8	124	4	AAU51146	797	17	70.8	165	6	ABM54594	ABM54594	Propionib
725	17	70.8	124	6	AAU47665	798	17	70.8	165	8	ADP31286	ADP31286	Human sec
726	17	70.8	127	8	ADP31622	799	17	70.8	165	8	ADP31515	ADP31515	Human sec
727	17	70.8	129	4	AAAG64067	800	17	70.8	165	8	ADP31174	ADP31174	Human sec
728	17	70.8	129	5	AAU96157	801	17	70.8	165	8	ADP31516	ADP31516	Human sec
729	17	70.8	129	6	AAU96129	802	17	70.8	165	8	ADP31041	ADP31041	Human sec
730	17	70.8	129	6	ABU97054	803	17	70.8	168	8	ADP31621	ADP31621	Human sec
731	17	70.8	130	4	AAAG63211	804	17	70.8	168	8	ADP30797	ADP30797	Human sec
732	17	70.8	130	4	AAAG64064	805	17	70.8	168	8	ADP30769	ADP30769	Human sec
733	17	70.8	130	4	AAU41366	806	17	70.8	168	8	ADP30771	ADP30771	Human sec
734	17	70.8	130	4	AAE09440	807	17	70.8	171	8	ADP30599	ADP30599	Human sec
735	17	70.8	130	5	AAU97612	808	17	70.8	173	3	AAAG54788	AAAG54788	Arabidops
736	17	70.8	130	5	AAU96153	809	17	70.8	175	7	ABO80411	ABO80411	Pseudomon
737	17	70.8	130	5	AAU10368	810	17	70.8	175	7	ABO69379	ABO69379	Pseudomon
738	17	70.8	130	5	AAU10366	811	17	70.8	177	7	ABO70036	ABO70036	Pseudomon
739	17	70.8	130	5	ABG74215	812	17	70.8	178	7	ABO74750	ABO74750	Pseudomon
740	17	70.8	130	6	ABG74206	813	17	70.8	178	7	ABO73284	ABO73284	Pseudomon
741	17	70.8	130	6	ABM37885	814	17	70.8	179	7	ABO75778	ABO75778	Pseudomon
742	17	70.8	130	6	ABP72422	815	17	70.8	180	8	ADP30820	ADP30820	Human sec
743	17	70.8	130	7	ADJ38121	816	17	70.8	180	8	ADP30825	ADP30825	Human sec
744	17	70.8	130	7	ABO81191	817	17	70.8	180	8	ADP30828	ADP30828	Human sec
745	17	70.8	130	8	ADP72483	818	17	70.8	180	8	ADP30821	ADP30821	Human sec
746	17	70.8	130	8	ADM86155	819	17	70.8	180	8	ADP30826	ADP30826	Human sec
747	17	70.8	130	9	ADM86175	820	17	70.8	180	8	ADP30827	ADP30827	Human sec
748	17	70.8	130	9	ABE12349	821	17	70.8	182	7	ABO78099	ABO78099	Pseudomon
749	17	70.8	130	5	AAU96155	822	17	70.8	182	7	ABO71666	ABO71666	Pseudomon
750	17	70.8	133	4	AAU49190	823	17	70.8	182	8	ADN99530	ADN99530	Novel hum
751	17	70.8	133	4	ABG04654	824	17	70.8	184	7	ABO77341	ABO77341	Pseudomon
752	17	70.8	133	6	ABM45709	825	17	70.8	185	8	ADP31109	ADP31109	Human sec
753	17	70.8	133	7	ABO82412	826	17	70.8	186	4	ABG08184	ABG08184	Novel hum
754	17	70.8	134	4	AAU58711	827	17	70.8	187	7	ABO80088	ABO80088	Pseudomon

828	17	70.8	190	6	ABM65810	901	17	70.8	249	3	AA60495	Agg60495	Arabidops
829	17	70.8	191	7	ABO76619	902	17	70.8	249	7	ABO83546	ABO83546	Pseudomon
830	17	70.8	192	8	ADP31335	903	17	70.8	250	6	ABP56993	ABP56993	E. tenell
831	17	70.8	192	4	ADP30575	904	17	70.8	250	7	ABO82689	ABO82689	Pseudomon
832	17	70.8	198	8	AAU47307	905	17	70.8	251	7	ADJ92161	ADJ92161	Human hai
833	17	70.8	198	6	ABM43826	906	17	70.8	251	7	ABO73872	ABO73872	Pseudomon
834	17	70.8	198	8	ADP30810	907	17	70.8	251	9	ABM91192	ABM91192	Human hai
835	17	70.8	198	8	ADP30811	908	17	70.8	254	8	ADP30739	ADP30739	Human sec
836	17	70.8	198	8	ADP30477	909	17	70.8	254	8	ADP31396	ADP31396	Human sec
837	17	70.8	198	8	ADP30481	910	17	70.8	255	7	ADJ92145	ADJ92145	Human hai
838	17	70.8	200	7	ABO71047	911	17	70.8	255	8	ADP30658	ADP30658	Human sec
839	17	70.8	201	7	ABO60877	912	17	70.8	256	4	ABG08337	ABG08337	Novel hum
840	17	70.8	203	7	ABD80007	913	17	70.8	256	6	ABU97138	ABU97138	Recombina
841	17	70.8	204	8	ADP31404	914	17	70.8	256	7	ABO75228	ABO75228	Pseudomon
842	17	70.8	206	7	ABO74022	915	17	70.8	256	8	ADN99494	ADN99494	Novel hum
843	17	70.8	208	4	AAU29506	916	17	70.8	257	5	ABP64573	ABP64573	Human ORF
844	17	70.8	208	5	ABG60794	917	17	70.8	258	8	ADP31084	ADP31084	Human sec
845	17	70.8	208	7	ABO80531	918	17	70.8	259	7	ABM8128	ABM8128	Rice abio
846	17	70.8	209	7	ABO77021	919	17	70.8	261	6	AAO27174	AAO27174	Pseudomon
847	17	70.8	210	8	ADP31563	920	17	70.8	263	3	AA60494	AA60494	Arabidops
848	17	70.8	210	8	ADP31562	921	17	70.8	263	7	ABO77554	ABO77554	Pseudomon
849	17	70.8	212	7	ABO67532	922	17	70.8	263	8	ABM81215	ABM81215	Tumour-as
850	17	70.8	215	7	ADP08255	923	17	70.8	264	8	ADP30788	ADP30788	Human sec
851	17	70.8	215	7	ABO72061	924	17	70.8	264	8	ADP31456	ADP31456	Human sec
852	17	70.8	215	9	ABM93776	925	17	70.8	264	8	ADP31412	ADP31412	Human sec
853	17	70.8	217	4	AAV97773	926	17	70.8	264	8	ADP31427	ADP31427	Human sec
854	17	70.8	218	7	ABO81158	927	17	70.8	266	8	ADP31564	ADP31564	Human sec
855	17	70.8	218	8	ADY23534	928	17	70.8	267	4	ABG27981	ABG27981	Novel hum
856	17	70.8	219	8	ADP31171	929	17	70.8	267	8	ADP31483	ADP31483	Human sec
857	17	70.8	219	8	ADY06385	930	17	70.8	270	8	ADP31217	ADP31217	Human sec
858	17	70.8	221	7	ADJ92147	931	17	70.8	270	8	ADP31321	ADP31321	Human sec
859	17	70.8	221	8	ADM87372	932	17	70.8	270	8	ADP31564	ADP31564	Human sec
860	17	70.8	222	5	ABM78572	933	17	70.8	270	9	ABE12365	ABE12365	Fusion pr
861	17	70.8	222	5	ABM78573	934	17	70.8	271	6	ABU6614	ABU6614	Protein e
862	17	70.8	223	3	ABM03862	935	17	70.8	271	6	ABU34590	ABU34590	Protein e
863	17	70.8	224	2	AAW22018	936	17	70.8	272	7	ADU92151	ADU92151	Human hai
864	17	70.8	224	2	ADW95939	937	17	70.8	272	6	ABU35898	ABU35898	Protein e
865	17	70.8	224	5	ADX15805	938	17	70.8	272	8	ADP12954	ADP12954	Protein e
866	17	70.8	225	5	ABM78570	939	17	70.8	273	8	ADP31236	ADP31236	Human sec
867	17	70.8	225	7	ABM88876	940	17	70.8	273	4	ABM94710	ABM94710	M. xanthu
868	17	70.8	225	8	ADP31501	941	17	70.8	274	4	AAU19418	AAU19418	Human dia
869	17	70.8	225	8	ADP30777	942	17	70.8	274	7	ABO72003	ABO72003	Pseudomon
870	17	70.8	227	7	ABO75458	943	17	70.8	274	8	ADP31679	ADP31679	Human sec
871	17	70.8	228	8	ADP30921	944	17	70.8	275	6	ABU34080	ABU34080	Protein e
872	17	70.8	231	5	ABP64726	945	17	70.8	276	7	ABO75093	ABO75093	Pseudomon
873	17	70.8	233	7	ABO75933	946	17	70.8	276	8	ADP31319	ADP31319	Human sec
874	17	70.8	235	7	ABO69554	947	17	70.8	276	8	ADP30584	ADP30584	Human sec
875	17	70.8	237	8	ADP31322	948	17	70.8	278	8	ADP31314	ADP31314	Human sec
876	17	70.8	238	3	ABE12367	949	17	70.8	279	8	ADP30846	ADP30846	Human sec
877	17	70.8	239	3	AAE21310	950	17	70.8	279	8	ADP31523	ADP31523	Human sec
878	17	70.8	239	8	ADY13038	951	17	70.8	279	8	ADP31503	ADP31503	Human sec
879	17	70.8	240	2	AAV22205	952	17	70.8	279	8	ADP31671	ADP31671	Human sec
880	17	70.8	240	8	ADP30737	953	17	70.8	279	8	ADP30605	ADP30605	Human sec
881	17	70.8	244	2	AAE44532	954	17	70.8	282	7	ADJ92143	ADJ92143	Human hai
882	17	70.8	244	2	AAW22985	955	17	70.8	282	7	ABO73521	ABO73521	Pseudomon
883	17	70.8	244	2	AAW51006	956	17	70.8	285	7	ABO73521	ABO73521	Pseudomon
884	17	70.8	244	3	ABE21323	957	17	70.8	288	8	ADP31463	ADP31463	Human sec
885	17	70.8	244	5	ABG65357	958	17	70.8	291	7	ABO68749	ABO68749	Pseudomon
886	17	70.8	244	6	AAE37572	959	17	70.8	293	6	ADAS5092	ADAS5092	Human pro
887	17	70.8	244	7	ADN80567	960	17	70.8	293	8	ADH09610	ADH09610	Human hos
888	17	70.8	244	7	ADN39212	961	17	70.8	294	8	ADP31076	ADP31076	Human sec
889	17	70.8	244	8	ADN39734	962	17	70.8	294	8	ADP31473	ADP31473	Human sec
890	17	70.8	244	8	ADN37158	963	17	70.8	297	8	ADP31192	ADP31192	Human sec
891	17	70.8	244	8	ADN04074	964	17	70.8	298	2	AAW43397	AAW43397	Human imi
892	17	70.8	244	8	ADN29289	965	17	70.8	302	3	AA607023	AA607023	Arabidops
893	17	70.8	244	8	ADN89076	966	17	70.8	304	8	ADP31659	ADP31659	Human sec
894	17	70.8	244	8	ADN72624	967	17	70.8	306	7	ADFO4021	ADFO4021	Bacterial
895	17	70.8	244	8	ADN72876	968	17	70.8	307	7	ADCO1287	ADCO1287	Enterococ
896	17	70.8	244	9	ADY67594	969	17	70.8	309	8	ADI42146	ADI42146	Plant tra
897	17	70.8	244	9	ADZ51357	970	17	70.8	310	4	ABBS5893	ABBS5893	Drosophil
898	17	70.8	245	7	ABO78304	971	17	70.8	311	4	ABB67853	ABB67853	Drosophil
899	17	70.8	246	8	ADP30619	972	17	70.8	311	4	ABG08405	ABG08405	Novel hum
900	17	70.8	246	9	ADY18495	973	17	70.8	312	1	AA650079	AA650079	T-cell an

```
974 17 70.8 312 1 AAP60471
975 17 70.8 312 2 AAR53145
976 17 70.8 312 3 AAB25555
977 17 70.8 312 4 ABG08404
978 17 70.8 312 8 ADP31505
979 17 70.8 312 8 ADP30476
980 17 70.8 313 7 ABM85966
981 17 70.8 315 5 AAE13348
982 17 70.8 315 6 ABR58684
983 17 70.8 315 8 ADP31685
984 17 70.8 316 3 AAY44862
985 17 70.8 316 9 ADY18921
986 17 70.8 317 8 ADP31615
987 17 70.8 318 7 ABG78573
988 17 70.8 318 8 ADP31066
989 17 70.8 320 4 AAG91631
990 17 70.8 325 3 AAB51703
991 17 70.8 326 3 AAG24798
992 17 70.8 336 8 ADP31320
993 17 70.8 341 7 ABG77462
994 17 70.8 343 7 AAG07022
995 17 70.8 345 3 AAG37538
996 17 70.8 345 8 ADP31204
997 17 70.8 346 7 ABO81210
998 17 70.8 346 7 ABM89803
999 17 70.8 346 8 ADP31355
1000 17 70.8 346 8 ADP31355
```

ALIGNMENTS

RESULT 1

ADQ13022 standard; peptide; 9 AA.

ADQ13022;

07-OCT-2004 (first entry)

Hepatitis B virus MHC class I restricted T-cell stimulating peptide 821.

HBV; MHC class I restricted T-cell stimulating peptide; surface peptide;

core peptide; polymerase peptide; immunogenic composition; vaccine;

cytotoxic T-lymphocyte response; CTL response.

Hepatitis B virus.

MO2004058807-A2.

15-JUL-2004.

09-DEC-2003; 2003WO-EP013948.

24-DEC-2002; 2002EP-00447276.

(ALGO-) ALGONOMICS NV.

Laetters I, Desmet J, Stegmann T, Castelein B;

WPI; 2004-525861/50.

New peptide comprising a major histocompatibility complex class I

restricted T-cell stimulating epitope of the hepatitis B virus (HBV)

surface, core and/or polymerase polypeptide, for preparation of an HBV

immunogenic composition.

Claim 5; SEQ ID NO 821; 108bp; English.

The invention comprises major histocompatibility complex (MHC) class I

restricted T-cell stimulating epitopes of the hepatitis B virus (HBV)

surface, core and polymerase proteins. The peptides of the invention are

useful for generating an HBV immunogenic composition (e.g. vaccine) that

CC induces a cytotoxic T-lymphocyte (CTL) response. The present amino acid
sequence represents an MHC class I restricted T-cell stimulating HBV
peptide of the invention.

SQ Sequence 9 AA;

Query Match 75.0%; Score 18; DB 8; Length 9;
Best Local Similarity 25.0%; Pred. No. 1.4e+06;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
DB 1 CASSSSSC 8

RESULT 2
ADQ13021 standard; peptide; 9 AA.

ADQ13021;

07-OCT-2004 (first entry)

Hepatitis B virus MHC class I restricted T-cell stimulating peptide 820.

HBV; MHC class I restricted T-cell stimulating peptide; surface peptide;

core peptide; polymerase peptide; immunogenic composition; vaccine;

cytotoxic T-lymphocyte response; CTL response.

Hepatitis B virus.

MO2004058807-A2.

15-JUL-2004.

09-DEC-2003; 2003WO-EP013948.

24-DEC-2002; 2002EP-00447276.

(ALGO-) ALGONOMICS NV.

Laetters I, Desmet J, Stegmann T, Castelein B;

WPI; 2004-525861/50.

New peptide comprising a major histocompatibility complex class I

restricted T-cell stimulating epitope of the hepatitis B virus (HBV)

surface, core and/or polymerase polypeptide, for preparation of an HBV

immunogenic composition.

Example 7; SEQ ID NO 820; 108bp; English.

The invention comprises major histocompatibility complex (MHC) class I

restricted T-cell stimulating epitopes of the hepatitis B virus (HBV)

surface, core and polymerase proteins. The peptides of the invention are

useful for generating an HBV immunogenic composition (e.g. vaccine) that

induces a cytotoxic T-lymphocyte (CTL) response. The present amino acid

sequence represents an MHC class I restricted T-cell stimulating HBV

peptide of the invention.

Sequence 9 AA;

Query Match 75.0%; Score 18; DB 8; Length 9;

Best Local Similarity 25.0%; Pred. No. 1.4e+06;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8

DB 2 CASSSSSC 9

RESULT 3
AAW65943

ID AAW65943 standard; peptide; 11 AA.
 XX AAW65943;
 AC
 XX 12-NOV-1998 (first entry)
 XX
 DE Molecule VI.
 XX
 KM Neurotrophin; cyclic; bicyclic; disulphide; nerve growth factor; NGF;
 KM BDNF; NT-3; conformation; promoter.
 XX
 OS Synthetic.
 OS Mammalia.
 XX CA2205045-A.
 XX 12-MAY-1998.
 PD
 XX 09-MAY-1997; 97CA-02205045.
 PF
 XX 12-NOV-1996; 96CA-02190296.
 PR
 XX (TOOH) UNIV QUEBENS KINGSTON.
 PA
 PI Riopelle RJ, Dory MI, Shamovsky IL, Weaver DF, Ross GM;
 XX WPI; 1998-457650/40.
 DR
 XX Inhibition of neurotrophin activity - using factor that interferes with
 PT sub-unit interaction.
 PS
 XX Disclosure; Page 13; 103pp; English.
 CC
 CC The invention relates to a method for reducing the biological activity of
 CC a multimeric protein having at least 2 promoters. It comprises (a)
 CC providing a factor that interacts with at least one portion of at least
 CC one of the promoters which associates with a portion of the other
 CC promoter in the absence of the factor; and (b) mixing the factor with the
 CC multimeric protein so that the factor interacts with the portion(s) and
 CC disrupts association of at least a portion of the promoters. The method
 CC and compounds are useful for inhibiting neurotrophin-mediated activities
 CC selected from neurotrophin receptor binding, neuron survival, neurite
 CC outgrowth and epileptic effects. The present sequence is shown in the
 CC specification
 CC
 SQ Sequence 11 AA;
 QY 1 CXXXXXXC 8
 Db 1 CAAAAAAC 8
 75.0%; Score 18; DB 2; Length 11;
 Best Local Similarity 25.0%; Pred. No. 2.3e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

RESULT 4
 AAW65944
 ID AAW65944 standard; peptide; 11 AA.
 XX
 AC AAW65944;
 XX
 DT 12-NOV-1998 (first entry)
 XX
 DE Molecule XV.
 XX
 KM Neurotrophin; cyclic; bicyclic; disulphide; nerve growth factor; NGF;
 KM BDNF; NT-3; conformation; promoter.
 XX
 OS Synthetic.
 OS Mammalia.
 XX CA2205045-A.
 PN

XX 12-MAY-1998.
 PD
 XX 09-MAY-1997; 97CA-02205045.
 PF
 XX 12-NOV-1996; 96CA-02190296.
 PR
 XX (TOOH) UNIV QUEBENS KINGSTON.
 PA
 PI Riopelle RJ, Dory MI, Shamovsky IL, Weaver DF, Ross GM;
 XX WPI; 1998-457650/40.
 DR
 XX Inhibition of neurotrophin activity - using factor that interferes with
 PT sub-unit interaction.
 PS
 XX Disclosure; Page 13; 103pp; English.
 CC
 CC The invention relates to a method for reducing the biological activity of
 CC a multimeric protein having at least 2 promoters. It comprises (a)
 CC providing a factor that interacts with at least one portion of at least
 CC one of the promoters which associates with a portion of the other
 CC promoter in the absence of the factor; and (b) mixing the factor with the
 CC multimeric protein so that the factor interacts with the portion(s) and
 CC disrupts association of at least a portion of the promoters. The method
 CC and compounds are useful for inhibiting neurotrophin-mediated activities
 CC selected from neurotrophin receptor binding, neuron survival, neurite
 CC outgrowth and epileptic effects. The present sequence is shown in the
 CC specification
 CC
 SQ Sequence 11 AA;
 QY 1 CXXXXXXC 8
 Db 4 CAAAAAAC 11
 75.0%; Score 18; DB 2; Length 11;
 Best Local Similarity 25.0%; Pred. No. 2.3e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

RESULT 5
 ADV23688
 ID ADV23688 standard; peptide; 15 AA.
 XX
 AC ADV23688;
 XX
 DT 10-MAR-2005 (first entry)
 DE HBV immunogenic peptide #489.
 XX
 KM Vaccine; virucide; antigen; autoimmune disease; infection;
 KM Immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;
 KM breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;
 KM pancreas tumor; stomach tumor; bladder tumor; kidney tumor;
 KM hodgekin's lymphoma.
 XX
 OS Hepatitis B virus.
 OS
 PN WO2004108753-A1.
 PD
 PD 16-DEC-2004.
 XX
 PF 10-JUN-2004; 2004MO-AUD000775.
 PR
 XX 10-JUN-2003; 2003AU-00902875.
 PR 25-MAR-2004; 2004AU-00901589.
 XX
 PA (UYME) UNIV MELBOURNE.
 PI Kent SJ;
 XX
 XX WPI; 2005-031657/03.
 DR

XX Use of at least one set of peptides in the preparation of a medicament
PT for modulating an immune response, and for treating cancer or yeast,
PT viral, bacterial, protozoal and mycoplasma infections.
XX
XX Disclosure; SEQ ID NO 2108; 645pp; English.
XX
XX The invention relates to the use of at least one set of peptides in the
CC preparation of a medicament for modulating an immune response, where
CC individual peptides of a respective set comprise different portions of an
CC amino acid sequence corresponding to a single polypeptide of interest and
CC display partial sequence identity or similarity to at least one other
CC peptide of the same set of peptides (i.e., they are overlapping). Also
CC included are an antigen-presenting cell which has been contacted with the
CC peptides above and thus presents the peptides, a population of such
CC antigen-presenting cells, a process for producing antigen-presenting
CC cells for modulating an immune response to a polypeptide of interest, a
CC method for producing antigen-specific lymphocytes, a composition
CC comprising at least one set of the peptides (and a carrier and/or
CC diluent), a method for modulating an immune response to a polypeptide of
CC interest comprising administering to a patient in need at least one set
CC of the peptides, a method for treatment and/or prophylaxis of a disease
CC or condition associated with the presence of a polypeptide of interest
CC and a composition of matter for modulating an immune response in a
CC subject to a target antigen. The polypeptide of interest is also a
CC disease- or condition-associated polypeptide that is a polypeptide
CC produced by a pathogenic organism or a cancer, and produced by a
CC pathogenic organism selected from yeast, viruses, bacteria, helminths,
CC protozoans and mycoplasmas. The disease- or condition-associated
CC polypeptide is produced by a cancer selected from melanoma, lung cancer,
CC breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic
CC cancer, stomach cancer, bladder cancer, kidney cancer, post transplant
CC lymphoproliferative disease (PTLD) or Hodgkin's lymphoma. The uncultured
CC antigen-presenting cells or their precursors are useful in the
CC preparation of a medicament for the treatment of a disease or condition
CC in a subject, which disease or condition is associated with the presence
CC of aberrant expression of a target antigen, where the antigen-presenting
CC cells or their precursors have not been subjected to activating
CC conditions but have been contacted with an antigen that corresponds to
CC the target antigen to express a processed or modified form of the antigen
CC for presentation to the subject's immune system. The present sequence is
CC one of a set of overlapping immunogenic peptides derived from a Hepatitis
CC B virus protein.
XX
SQ Sequence 15 AA;
Query Match 75.0%; Score 18; DB 9; Length 15;
Best Local Similarity 25.0%; Pred. No. 2.4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CXXXXXXC 8
DB 4 CASSSSSC 11
RESULT 6
ADV23687
ID ADV23687 standard; peptide; 15 AA.
XX
XX ADV23687;
XX
XX 10-MAR-2005 (first entry)
XX
XX HBV immunogenic peptide #488.
XX
XX Vaccine; virucide; antigen; autoimmune disease; infection;
XX immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;
XX breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;
XX pancreas tumor; stomach tumor; bladder tumor; kidney tumor;
XX hodgekin's lymphoma.
XX
XX Hepatitis B virus.
XX

PN WO2004108753-A1.
XX
XX 16-DEC-2004.
XX
XX 10-JUN-2004; 2004WO-AU000775.
XX
XX 10-JUN-2003; 2003AU-00902875.
XX
XX 25-MAR-2004; 2004AU-00901589.
XX
XX (UTME) UNIV MELBOURNE.
XX
XX Kent SJ;
XX
XX WPI; 2005-031657/03.
XX
XX
XX Use of at least one set of peptides in the preparation of a medicament
PT for modulating an immune response, and for treating cancer or yeast,
PT viral, bacterial, protozoal and mycoplasma infections.
XX
XX Disclosure; SEQ ID NO 2107; 645pp; English.
XX
XX The invention relates to the use of at least one set of peptides in the
CC preparation of a medicament for modulating an immune response, where
CC individual peptides of a respective set comprise different portions of an
CC amino acid sequence corresponding to a single polypeptide of interest and
CC display partial sequence identity or similarity to at least one other
CC peptide of the same set of peptides (i.e., they are overlapping). Also
CC included are an antigen-presenting cell which has been contacted with the
CC peptides above and thus presents the peptides, a population of such
CC antigen-presenting cells, a process for producing antigen-presenting
CC cells for modulating an immune response to a polypeptide of interest, a
CC method for producing antigen-specific lymphocytes, a composition
CC comprising at least one set of the peptides (and a carrier and/or
CC diluent), a method for modulating an immune response to a polypeptide of
CC interest comprising administering to a patient in need at least one set
CC of the peptides, a method for treatment and/or prophylaxis of a disease
CC or condition associated with the presence of a polypeptide of interest
CC and a composition of matter for modulating an immune response in a
CC subject to a target antigen. The polypeptide of interest is also a
CC disease- or condition-associated polypeptide that is a polypeptide
CC produced by a pathogenic organism or a cancer, and produced by a
CC pathogenic organism selected from yeast, viruses, bacteria, helminths,
CC protozoans and mycoplasmas. The disease- or condition-associated
CC polypeptide is produced by a cancer selected from melanoma, lung cancer,
CC breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic
CC cancer, stomach cancer, bladder cancer, kidney cancer, post transplant
CC lymphoproliferative disease (PTLD) or Hodgkin's lymphoma. The uncultured
CC antigen-presenting cells or their precursors are useful in the
CC preparation of a medicament for the treatment of a disease or condition
CC in a subject, which disease or condition is associated with the presence
CC of aberrant expression of a target antigen, where the antigen-presenting
CC cells or their precursors have not been subjected to activating
CC conditions but have been contacted with an antigen that corresponds to
CC the target antigen to express a processed or modified form of the antigen
CC for presentation to the subject's immune system. The present sequence is
CC one of a set of overlapping immunogenic peptides derived from a Hepatitis
CC B virus protein.
XX
SQ Sequence 15 AA;
Query Match 75.0%; Score 18; DB 9; Length 15;
Best Local Similarity 25.0%; Pred. No. 2.4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CXXXXXXC 8
DB 8 CASSSSSC 15
RESULT 7
ABB42273
ID ABB42273 standard; peptide; 24 AA.
XX


```
AC ABB42273;
XX
DT 04-FEB-2002 (first entry)
DE Peptide #9779 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
PN WO200157277-A2.
PD 09-AUG-2001.
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632356.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
PA (MOLE-) MOLECULAR DYNAMICS INC.
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human fetal liver.
PS Claim 27; SEQ ID NO 34908; 639bp + Sequence Listing; English.
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 24 AA;
Query Match 75.0%; Score 18; DB 4; Length 24;
Best Local Similarity 25.0%; Pred. No. 2.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0.
OY 1 CXXXXXXC 8
Db | |
12 CTTS TTC 19
RESULT 8
ID AAM36081
AAC AAM36081 standard; protein; 24 AA.
AC AAM36081;
XX
DT 17-OCT-2001 (first entry)
DE Peptide #10118 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
```

[illegible]

DR WPI; 2001-488900/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 36279; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
SQ Sequence 24 AA;

Query Match 75.0%; Score 18; DB 4; Length 24;
Best Local Similarity 25.0%; Pred. No. 2.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
DB 12 CTTSTTTC 19

RESULT 10
AAM63159
ID AAM63159 standard; protein; 24 AA.
XX
AC AAM63159;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35264.
XX
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
KM Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 35264; 650bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX

SQ Sequence 24 AA;

Query Match 75.0%; Score 18; DB 4; Length 24;
Best Local Similarity 25.0%; Pred. No. 2.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
DB 12 CTTSTTTC 19

RESULT 11
ABG57702
ID ABG57702 standard; peptide; 24 AA.
XX
AC ABG57702;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver peptide, SEQ ID NO 36350.
XX
XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KM hypercholesterolaemia; coronary heart disease.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PF Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 27; SEQ ID NO 36350; 658bp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59530 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 24 AA;

Query Match 75.0%; Score 18; DB 4; Length 24;
Best Local Similarity 25.0%; Pred. No. 2.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8

Db 12 CTTSTTTC 19

RESULT 12
ABB68695
ID ABB68695 standard; protein; 69 AA.
XX
AC ABB68695;
XX
DT 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 32877.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX Drosophila melanogaster.
XX OS
PN WO300171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-6556860/75.
XX N-PSDB; ABL12798.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX
PS Disclosure; SEQ ID NO 32877; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins (AB57737-
XX AB572072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pcc_sequences
XX
SQ Sequence 69 AA;
XX
Query Match 75.0%; Score 18; DB 4; Length 69;
Best Local Similarity 25.0%; Pred. No. 3.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CXXXXXXC 8
Db 9 CSSAASC 16

RESULT 13
ADP30715
ID ADP30715 standard; protein; 69 AA.
XX
AC ADP30715;
XX
DT 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #1482.
XX

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411049P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411101P.
XX 17-SEP-2002; 2002US-0411111P.
XX 18-APR-2003; 2003US-0463700P.
XX 18-APR-2003; 2003US-0463708P.
XX 18-APR-2003; 2003US-0463716P.
XX 18-APR-2003; 2003US-0463732P.
XX 02-MAY-2003; 2003US-0467189P.
XX 02-MAY-2003; 2003US-0467201P.
XX 02-MAY-2003; 2003US-0467203P.
XX 02-MAY-2003; 2003US-0467230P.
XX 19-MAY-2003; 2003US-0471306P.
XX 19-MAY-2003; 2003US-0471336P.
XX 22-MAY-2003; 2003US-0472420P.
XX 22-MAY-2003; 2003US-0472430P.
XX 09-JUN-2003; 2003US-0476609P.
XX 09-JUN-2003; 2003US-0476641P.
XX 08-JUL-2003; 2003US-0485218P.
XX 08-JUL-2003; 2003US-0485223P.
XX 08-JUL-2003; 2003US-0485224P.
XX 08-JUL-2003; 2003US-0485325P.
XX 14-JUL-2003; 2003US-0486446P.
XX 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
PS Claim 1; SEQ ID NO 2713; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cyostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMED and is not in the specification.
XX
SQ Sequence 69 AA;

Query Match 75.0%; Score 18; DB 8; Length 69;
Best Local Similarity 25.0%; Pred. No. 3.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 18 CTTTAAC 25

RESULT 14
ABB39058
ID ABB39058 standard; peptide; 91 AA.
XX
XX ABB39058;
XX
XX 04-FEB-2002 (first entry)
XX
XX Peptide #6564 encoded by human foetal liver single exon probe.
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
XX Homo sapiens.
XX
XX WO200157277-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000669.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human foetal liver.
XX
XX
PS Claim 27; SEQ ID NO 31693; 639pp + Sequence listing; English.
XX

XX The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human foetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 91 AA;
Query Match 75.0%; Score 18; DB 4; Length 91;
Best Local Similarity 25.0%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 29 CSSSSSSC 36

RESULT 15
AAM32549
ID AAM32549 standard; protein; 91 AA.
XX
XX AAM32549;
XX

DT 17-OCT-2001 (first entry)
XX
XX Peptide #6586 encoded by probe for measuring placental gene expression.
XX

XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX

XX Homo sapiens.
XX

XX WO200157272-A2.
XX

XX 09-AUG-2001.
XX

XX 30-JAN-2001; 2001WO-US000663.
XX

XX 04-FEB-2000; 2000US-0180312P.
XX

XX 26-MAY-2000; 2000US-0207456P.
XX

XX 30-JUN-2000; 2000US-00608408.
XX

XX 03-AUG-2000; 2000US-00632366.
XX

XX 21-SEP-2000; 2000US-0234687P.
XX

XX 27-SEP-2000; 2000US-0236359P.
XX

XX 04-OCT-2000; 2000GB-00024263.
XX

XX (MOLE-) MOLECULAR DYNAMICS INC.
XX

XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX

XX WPI; 2001-488897/53.
XX

XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX

XX Claim 27; SEQ ID NO 32818; 654pp; English.
XX

XX The present invention relates to single exon nucleic acid probes (SENP:
CC see A113315-A157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders

```
XX
SQ Sequence 91 AA;
Query Match 75.0%; Score 18; DB 4; Length 91;
Best Local Similarity 25.0%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Oy 1 CXXXXXXC 8
Db 29 CSSSSSSC 36

RESULT 16
AAM72290
ID AAM72290 standard; protein; 91 AA.
XX
AC AAM72290;
XX
XX 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 32596.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 32596; 658bp + Sequence listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
SQ Sequence 91 AA;
Query Match 75.0%; Score 18; DB 4; Length 91;
Best Local Similarity 25.0%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Oy 1 CXXXXXXC 8
Db 29 CSSSSSSC 36

RESULT 17
AAM59710
ID AAM59710 standard; protein; 91 AA.
```

```
XX
AC AAM59710;
XX
XX 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 31815.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 31815; 650bp + Sequence listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX
SQ Sequence 91 AA;
Query Match 75.0%; Score 18; DB 4; Length 91;
Best Local Similarity 25.0%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Oy 1 CXXXXXXC 8
Db 29 CSSSSSSC 36

RESULT 18
ABG53976
ID ABG53976 standard; peptide; 91 AA.
XX
AC ABG53976;
XX
XX 25-FEB-2003 (first entry)
XX
DE Human liver peptide; SEQ ID NO 32624.
XX
KW Human; liver; cirrhosis; hyperlipoproteinemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
```

PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000664.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488898/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human adult liver.
 XX
 PS Claim 27; SEQ ID NO 32624; 658bp; English.
 XX
 CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult liver.
 CC (I) may be used for predicting, measuring and displaying gene expression
 CC in samples derived from human adult liver. The genes identified may be
 CC involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
 CC associated with coronary heart disease. ABG47348-ABG59930 represent human
 CC liver single exon encoded peptides of the invention. Note: The sequence
 CC information for this patent does not appear in the printed specification
 CC but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX Sequence 91 AA;
 SQ
 Query Match 75.0%; Score 18; DB 4; Length 91;
 Best Local Similarity 25.0%; Pred. No. 3.5e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CXXXXXXC 8
 Db 29 CSSSSSSC 36
 Db
 RESULT 19
 ABG42105
 ID ABG42105 standard; peptide: 91 AA.
 XX
 AC ABG42105;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 31770.
 XX
 KW Human: single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX

PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000665.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.
 XX
 PS Claim 27; SEQ ID NO 31770; 634bp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC; the novel set of probes which hybridise at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC collected from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a peptide/protein encoded by a single exon probe of
 CC the invention. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 XX Sequence 91 AA;
 SQ
 Query Match 75.0%; Score 18; DB 5; Length 91;
 Best Local Similarity 25.0%; Pred. No. 3.5e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CXXXXXXC 8
 Db 29 CSSSSSSC 36
 Db

RESULT 20
ADP30859
ID ADP30859 standard; protein; 93 AA.
XX AC ADP30859;
XX DT 12-AUG-2004 (first entry)
XX DE Human secreted protein SEQ ID #1626.
XX DE Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX KW cancer; inflammatory; immune; human secreted protein.
XX OS Homo sapiens.
XX PN WO2004035732-A2.
XX PD 29-APR-2004.
XX PF 28-AUG-2003; 2003WO-US026780.
XX PR 29-AUG-2002; 2002US-0406576P.
XX PR 29-AUG-2002; 2002US-0406579P.
XX PR 29-AUG-2002; 2002US-0406585P.
XX PR 29-AUG-2002; 2002US-0406588P.
XX PR 29-AUG-2002; 2002US-0406608P.
XX PR 29-AUG-2002; 2002US-0406611P.
XX PR 29-AUG-2002; 2002US-0406612P.
XX PR 29-AUG-2002; 2002US-0406616P.
XX PR 29-AUG-2002; 2002US-0406640P.
XX PR 29-AUG-2002; 2002US-0406642P.
XX PR 29-AUG-2002; 2002US-0406653P.
XX PR 29-AUG-2002; 2002US-0406655P.
XX PR 29-AUG-2002; 2002US-0406666P.
XX PR 17-SEP-2002; 2002US-0410946P.
XX PR 17-SEP-2002; 2002US-0410947P.
XX PR 17-SEP-2002; 2002US-0410948P.
XX PR 17-SEP-2002; 2002US-0410949P.
XX PR 17-SEP-2002; 2002US-0410953P.
XX PR 17-SEP-2002; 2002US-0410957P.
XX PR 17-SEP-2002; 2002US-0410958P.
XX PR 17-SEP-2002; 2002US-0410959P.
XX PR 17-SEP-2002; 2002US-0410960P.
XX PR 17-SEP-2002; 2002US-0410961P.
XX PR 17-SEP-2002; 2002US-0410962P.
XX PR 17-SEP-2002; 2002US-0411019P.
XX PR 17-SEP-2002; 2002US-0411023P.
XX PR 17-SEP-2002; 2002US-0411024P.
XX PR 17-SEP-2002; 2002US-0411032P.
XX PR 17-SEP-2002; 2002US-0411035P.
XX PR 17-SEP-2002; 2002US-0411037P.
XX PR 17-SEP-2002; 2002US-0411041P.
XX PR 17-SEP-2002; 2002US-0411045P.
XX PR 17-SEP-2002; 2002US-0411046P.
XX PR 17-SEP-2002; 2002US-0411048P.
XX PR 17-SEP-2002; 2002US-0411052P.
XX PR 17-SEP-2002; 2002US-0411055P.
XX PR 17-SEP-2002; 2002US-0411073P.
XX PR 17-SEP-2002; 2002US-0411082P.
XX PR 17-SEP-2002; 2002US-0411101P.
XX PR 17-SEP-2002; 2002US-0411111P.
XX PR 18-APR-2003; 2003US-0463700P.
XX PR 18-APR-2003; 2003US-0463708P.
XX PR 18-APR-2003; 2003US-0463716P.
XX PR 18-APR-2003; 2003US-0463732P.
XX PR 02-MAY-2003; 2003US-0467199P.
XX PR 02-MAY-2003; 2003US-0467201P.
XX PR 02-MAY-2003; 2003US-0467203P.
XX PR 02-MAY-2003; 2003US-0467230P.
XX PR 19-MAY-2003; 2003US-0471306P.
XX PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485229P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Haisan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX DR WPI; 2004-348438/32.
XX PR New nucleic acid molecule for diagnosing, preventing or treating diseases
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX PT genetic, bacterial and viral diseases.
XX PS Claim 1; SEQ ID NO 2857; 428bp; English.
XX CC The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cyostatic,
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPOMEB and is not in the specification.
XX SQ Sequence 93 AA;
Query Match 75.0%; Score 18; DB 8; Length 93;
Best Local Similarity 25.0%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
CY 1 CXXXXXXC 8
DB 85 CATATAC 92
RESULT 21
ABO83167
ID ABO83167 standard; protein; 99 AA.
XX AC ABO83167;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polypeptide #15342.
XX DE Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX KW Pseudomonas aeruginosa.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.
PA Rubenstein MJ, Nolling J, Deloughery C, Bush D;
PI WPI; 2003-615309/58.
XX N-PSDB; ABD16738.
DR Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
PS Disclosure; SEQ ID NO 31913; 455bp; English.
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using bioclip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC specification data for this patent did not form part of the printed
CC seqdata.uspto.gov/sequence.html
XX Sequence 99 AA;
QY 1 CXXXXXXC 8
Db 24 CSTRASAC 31
Query Match 75.0%; Score 18; DB 7; Length 99;
Best Local Similarity 25.0%; Pred. No. 3.6e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
RESULT 22
ADP30709
XX ADP30709 standard; protein; 99 AA.
AC ADP30709;
XX 12-AUG-2004 (first entry)
DE Human secreted protein SEQ ID #1476.
XX Cytosol; Anticarcinogenic; Immunosuppressive; Antibacterial; Virucide;
KW cancer; Inflammatory; Immune; human secreted protein.
XX Homo sapiens.
OS WO2004035732-A2.
PN 29-APR-2004.
XX 29-APR-2004.
PD 28-AUG-2003; 2003WO-US026780.
PF 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
XX 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
XX 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
XX 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
XX 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
XX 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471316P.
XX 22-MAY-2003; 2003US-0472430P.
PR 22-MAY-2003; 2003US-0472430P.
XX 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
XX 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
XX 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485246P.
XX 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
XX 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
XX 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
XX 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Helsenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX Claim 1; SEQ ID NO 2707; 428bp; English.
XX

CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytosolic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer, inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
SQ Sequence 99 AA;

Query Match 75.0%; Score 18; DB 8; Length 99;
Best Local Similarity 25.0%; Pred. No. 3.6e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 77 CAAATATC 84

RESULT 23

ID ABO74878 standard; protein; 107 AA.

XX ABO74878;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #7053.

KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

PA (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

DR N-PSDB; ABD08449.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 23624; 455pp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX

SQ Sequence 107 AA;

Query Match 75.0%; Score 18; DB 7; Length 107;
Best Local Similarity 25.0%; Pred. No. 3.6e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 29 CSAASSAC 36

RESULT 24

ID ABO73394 standard; protein; 115 AA.

XX ABO73394;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #5569.

KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

PA (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

DR N-PSDB; ABD06965.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 22140; 455pp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX

SQ Sequence 115 AA;

Query Match 75.0%; Score 18; DB 7; Length 115;
Best Local Similarity 25.0%; Pred. No. 3.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 83 CSAATSTC 90

PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493572P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
DR
XX
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
XX Claim 1; SEQ ID NO 2701; 428bp; English.
PS
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic.
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
XX Sequence 135 AA;
SQ

Query Match 75.0%; Score 18; DB 8; Length 135;
Best Local Similarity 25.0%; Pred. No. 3.8e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CXXXXXXC 8
Db 105 CTTTATC 112

RESULT 27
ABO75730
ID ABO75730 standard; protein; 144 AA.
XX
XX ABO75730;
AC
XX
XX 29-JUL-2004 (first entry)
DT
XX
XX Pseudomonas aeruginosa polypeptide #7905.
DB
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
KW
XX Pseudomonas aeruginosa.
OS
XX
XX US651795-B1.
PN
XX
XX 22-APR-2003.
PD
XX
XX 18-FEB-1999; 99US-002522991.
PF
XX
XX 18-FEB-1998; 98US-0074788P.
PR
XX 27-JUL-1998; 98US-0094190P.
PR
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI
XX WPI; 2003-615309/58.
PI
XX N-PSDB; ABD09301.
DR
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of

PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 24476; 455bp; English.
PS
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO64336 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX Sequence 144 AA;
SQ

Query Match 75.0%; Score 18; DB 7; Length 144;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CXXXXXXC 8
Db 36 CSATSAC 43

RESULT 28
ADP31453
ID ADP31453 standard; protein; 144 AA.
XX
XX ADP31453;
AC
XX
XX 12-AUG-2004 (first entry)
DT
XX
XX Human secreted protein SEQ ID #2220.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
KW
XX Homo sapiens.
OS
XX
XX WO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PF
XX
XX 29-AUG-2002; 2002US-0406576P.
PR
XX 29-AUG-2002; 2002US-0406579P.
PR
XX 29-AUG-2002; 2002US-0406585P.
PR
XX 29-AUG-2002; 2002US-0406588P.
PR
XX 29-AUG-2002; 2002US-0406608P.
PR
XX 29-AUG-2002; 2002US-0406611P.
PR
XX 29-AUG-2002; 2002US-0406612P.
PR
XX 29-AUG-2002; 2002US-0406616P.
PR
XX 29-AUG-2002; 2002US-0406642P.
PR
XX 29-AUG-2002; 2002US-0406646P.
PR
XX 29-AUG-2002; 2002US-0406653P.
PR
XX 29-AUG-2002; 2002US-0406655P.
PR
XX 29-AUG-2002; 2002US-0406666P.
PR
XX 17-SEP-2002; 2002US-0410946P.
PR
XX 17-SEP-2002; 2002US-0410947P.
PR
XX 17-SEP-2002; 2002US-0410948P.
PR
XX 17-SEP-2002; 2002US-0410949P.
PR
XX 17-SEP-2002; 2002US-0410953P.
PR
XX 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486466P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Heestir K, Beaurang PA, Behrens D;
PI Haldenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3451; 428pp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytosolic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMES and is not in the specification.
XX
XX Sequence 144 AA;

Query Match 75.0%; Score 18; DB 8; Length 144;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CXXXXXXC 8
Db 20 CTATATAC 27
RESULT 29
ADX90450
ID ADX90450 standard; protein; 144 AA.
XX
XX ADX90450;
AC
XX 21-APR-2005 (first entry)
DT
XX
XX Plant full length insert polypeptide seqid 53114.
DE
XX
XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
XX Unidentified.
OS
XX
XX US2004034888-A1.
PN
XX
XX 19-FEB-2004.
PD
XX
XX 28-APR-2003; 2003US-00425114.
PF
XX
XX 06-MAY-1999; 99US-00304517.
PR
XX
XX 05-NOV-2001; 2001US-00985678.
PR
XX
XX (LITU/) LITU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABAS/) TABASKA J E.
PA (CAOV/) CAO Y.
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
PI
XX
XX WPI; 2004-180133/17.
DR
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX Claim 1; SEQ ID NO 53114; 15pp; English.
PS
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or

CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 144 AA;
Query Match 75.0%; Score 18; DB 8; Length 144;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CXXXXXXC 8
DB 4 CTTSSAAC 11
RESULT 30
ABU45351
ID ABU45351 standard; protein; 150 AA.
XX
AC ABU45351;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #30878.
XX
DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
KW Salmomella paratyphi.
XX
OS Salmomella paratyphi.
XX
FN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX
PR 06-SEP-2001; 2001US-00948993.
XX
PR 25-OCT-2001; 2001US-0342923P.
XX
PR 08-FEB-2002; 2002US-00072851.
XX
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI: 2003-029926/02.
XX
DR N-PSDB; ACA49221.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 73275; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 623 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 150 AA;
Query Match 75.0%; Score 18; DB 6; Length 150;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CXXXXXXC 8
DB 16 CATTASCC 23
RESULT 31
AAB41043
ID AAB41043 standard; protein; 153 AA.
XX
AC AAB41043;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORF807 polypeptide sequence SEQ ID NO:1614.
XX
XX Human; open reading frame; ORF; detection; cytostatic; hepatotropic;
XX vulnery; antiparasitic; antiparkinsonian; nootropic; neuroprotective;
XX anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX hypotensive; dermatological; immunosuppressive; antiinflammatory;
XX antiviral; antibacterial; antifungal; antineumatic; antithyroid;
XX antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX cholesterol ester storage; systemic lupus erythematosus; infection;
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX bone damage; cartilage damage; antiinflammatory disease; coagulation;
XX thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
FN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US008621.
XX
PR 31-MAR-1999; 99US-0127607P.
XX
PR 02-APR-1999; 99US-0127636P.
XX
PR 05-APR-1999; 99US-0127728P.
XX
PR 30-MAR-2000; 2000US-00540763.
XX
PA (CIRRA-) CIRAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI: 2000-602362/57.
XX
DR N-PSDB; AAC75252.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.

XX Claim 11; Page 1302; 5507pp; English.
 PS
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatocytic; valnery;
 CC antiproliferative; antiparkinsonian; neurotrophic; neuroprotective; osteopathic;
 CC anticonvulsant; antichrystic; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatologic; immunosuppressive; antiinflammatory; antibacterial;
 CC antiviral; antifungal; antirheumatic; antichryoid; and antinaemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 CC enhance coagulation, to inhibit thrombosis; and as a contraceptive
 CC
 XX
 SQ Sequence 153 AA;
 Query Match 75.0%; Score 18; DB 3; Length 153;
 Best Local Similarity 25.0%; Pred. No. 3.9e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CXXXXXXC 8
 Db 78 CSSSASC 85
 RESULT 32
 ABP09251
 ID ABP09251 standard; protein; 153 AA.
 XX
 AC ABP09251;
 DT 25-JUN-2002 (first entry)
 XX
 DE Human ORFX protein sequence SEQ ID NO:18484.
 XX
 KW Human, open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.
 KW
 OS Homo sapiens.
 XX
 PN WO200192523-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 29-MAY-2001; 2001WO-US010836.
 XX
 PR 30-MAY-2000; 2000US-0206132P.
 XX
 PR 29-AUG-2000; 2000US-0228716P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimkets RA, Leach MD;
 XX
 DR WPI; 2002-106308/14.
 XX
 DR N-PSDB; ABN25003.
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,

PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders.
 XX
 PS Disclosure; SEQ ID NO 18484; 1037pp; English.
 XX
 CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN7252 encode the human ORFX
 CC proteins given in ABP0010 to ABP1500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 153 AA;
 Query Match 75.0%; Score 18; DB 5; Length 153;
 Best Local Similarity 25.0%; Pred. No. 3.9e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CXXXXXXC 8
 Db 78 CSSSASC 85
 RESULT 33
 ABO74421
 ID ABO74421 standard; protein; 154 AA.
 XX
 AC ABO74421;
 DT 29-JUL-2004 (first entry)
 XX
 DE Pseudomonas aeruginosa polypeptide #6596.
 XX
 KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 KW
 OS Pseudomonas aeruginosa.
 XX
 PN US651795-B1.
 XX
 PD 22-APR-2003.
 XX
 PF 18-FEB-1999; 99US-00252991.
 XX
 PR 18-FEB-1998; 98US-0074788P.
 XX
 PR 27-JUL-1998; 98US-0094190P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 XX
 DR WPI; 2003-615309/58.
 XX
 DR N-PSDB; ABD07992.
 XX
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of

PT pathological conditions resulting from bacterial infection.
XX
XX Disclousure; SEQ ID NO 21167; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX Sequence 154 AA;
SQ
Query Match 75.0%; Score 18; DB 7; Length 154;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CXXXXXXC 8
Db 120 CAAAASAC 127
RESULT 34
ABO77836 ID ABO77836 standard; protein, 156 AA.
XX
XX ABO77836;
XX
DT 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polypeptide #10011.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
XX
XX US6551795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX
XX 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
XX WPI; 2003-615309/58.
XX
XX N-PSDB; ABD11407.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX
XX Disclousure; SEQ ID NO 26582; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,

CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX Sequence 156 AA;
SQ
Query Match 75.0%; Score 18; DB 7; Length 156;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CXXXXXXC 8
Db 40 CSASSAC 47
RESULT 35
ADP30761 ID ADP30761 standard; protein, 159 AA.
XX
XX ADP30761;
XX
DT 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #1528.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX
XX 29-AUG-2002; 2002US-0406579P.
XX
XX 29-AUG-2002; 2002US-0406585P.
XX
XX 29-AUG-2002; 2002US-0406588P.
XX
XX 29-AUG-2002; 2002US-0406608P.
XX
XX 29-AUG-2002; 2002US-0406611P.
XX
XX 29-AUG-2002; 2002US-0406616P.
XX
XX 29-AUG-2002; 2002US-0406640P.
XX
XX 29-AUG-2002; 2002US-0406642P.
XX
XX 29-AUG-2002; 2002US-0406646P.
XX
XX 29-AUG-2002; 2002US-0406653P.
XX
XX 29-AUG-2002; 2002US-0406655P.
XX
XX 29-AUG-2002; 2002US-0406666P.
XX
XX 17-SEP-2002; 2002US-0410946P.
XX
XX 17-SEP-2002; 2002US-0410947P.
XX
XX 17-SEP-2002; 2002US-0410948P.
XX
XX 17-SEP-2002; 2002US-0410949P.
XX
XX 17-SEP-2002; 2002US-0410953P.
XX
XX 17-SEP-2002; 2002US-0410957P.
XX
XX 17-SEP-2002; 2002US-0410958P.
XX
XX 17-SEP-2002; 2002US-0410959P.
XX
XX 17-SEP-2002; 2002US-0410960P.
XX
XX 17-SEP-2002; 2002US-0410961P.
XX
XX 17-SEP-2002; 2002US-0410962P.
XX
XX 17-SEP-2002; 2002US-0411019P.
XX
XX 17-SEP-2002; 2002US-0411022P.
XX
XX 17-SEP-2002; 2002US-0411023P.
XX
XX 17-SEP-2002; 2002US-0411024P.

PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486961P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493577P.
PR 08-AUG-2003; 2003US-0493577P.
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Heetir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang WM, Kothakota S, Haishan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
DR New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2759; 428bp; English.
PS
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cyostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
XX Sequence 159 AA;

Query Match 75.0%; Score 18; DB 8; Length 159;
Best Local Similarity 25.0%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
DB 125 CAAATATC 132

RESULT 36
ADP30758
ID ADP30758 standard; protein; 162 AA.
XX
XX ADP30758;
XX
XX 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #1525.
XX
XX Cyostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411101P.
XX 17-SEP-2002; 2002US-0411111P.
XX 18-APR-2003; 2003US-0463700P.
XX 18-APR-2003; 2003US-0463708P.
XX 18-APR-2003; 2003US-0463716P.
XX 18-APR-2003; 2003US-0463732P.
XX 02-MAY-2003; 2003US-0467199P.
XX 02-MAY-2003; 2003US-0467201P.
XX 02-MAY-2003; 2003US-0467203P.
XX 02-MAY-2003; 2003US-0467230P.
XX 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-047609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang WM, Kochakota S, Hsiehan L, Linnemann T;
PI Pierce K, Wang Y, Wong USP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 2756; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
SQ Sequence 162 AA;
QY
Db 1 CXXXXXXC 8
53 CAATTATC 60
Query Match 75.0%; Score 18; DB 8; Length 162;
Best Local Similarity 25.0%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
RESULT 37
ADP30756
ID ADP30756 standard; protein; 162 AA.
XX
AC ADP30756;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1523.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-USO26780.
XX

PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463709P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
XX PI Halenbeck RF, Huang WM, Kochakota S, Hsiehan L, Linnemann T;
PI

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2754; 428bp; English.
PS
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
XX Sequence 162 AA;
SQ
Query Match 75.0%; Score 18; DB 8; Length 162;
Best Local Similarity 25.0%; Pred. No. 4e+02; 6; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 0;
QY 1 CXXXXXXC 8
Db 53 CAATTATC 60
RESULT 38
ADP30757
ID ADP30757 standard; protein; 162 AA.
XX
AC ADP30757;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1524.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; Inflammatory; Immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406665P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411057P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471366P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476611P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang WM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2755; 428bp; English.
PS
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
XX Sequence 162 AA;
SQ
Query Match 75.0%; Score 18; DB 8; Length 162;

Best Local Similarity 25.0%; Pred. No. 4e+02; Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
Db 53 CATTATTC 60

RESULT 39

ID ADP30813 standard; protein, 162 AA.

AC ADP30813;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #1580.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

OS WO2004035732-A2.

PN 29-APR-2004.

PD 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

PR 29-AUG-2002; 2002US-0406579P.

PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.

PR 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406616P.

PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.

PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.

PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.

PR 17-SEP-2002; 2002US-0410961P.

PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.

PR 17-SEP-2002; 2002US-0411022P.

PR 17-SEP-2002; 2002US-0411023P.

PR 17-SEP-2002; 2002US-0411032P.

PR 17-SEP-2002; 2002US-0411035P.

PR 17-SEP-2002; 2002US-0411037P.

PR 17-SEP-2002; 2002US-0411041P.

PR 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.

PR 17-SEP-2002; 2002US-0411048P.

PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.

PR 17-SEP-2002; 2002US-0411101P.

PR 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463700P.

PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463716P.

PR 18-APR-2003; 2003US-0463732P.

PR 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.

PR 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467203P.

PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.

PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.

PR 09-JUN-2003; 2003US-0476641P.

PR 08-JUL-2003; 2003US-0485218P.

PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485224P.

PR 08-JUL-2003; 2003US-0485325P.

PR 14-JUL-2003; 2003US-0486467P.

PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486960P.

PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.

PR 08-AUG-2003; 2003US-0493577P.

XX (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;

PI Halenbeck RF, Huang WM, Kothakota S, Halshan L, Linnemann T;

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

XX WPI, 2004-348436/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases

PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,

PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 2811; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule

CC encoding a polypeptide which is believed to be cytosstatic, the

CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and

CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC immune, metabolic, genetic, bacterial and viral diseases. The present

CC sequence represents a human secreted protein. The present sequence is

CC available on WIPOMB and is not in the specification.

SQ Sequence 162 AA;

Query Match 75.0%; Score 18; DB 8; Length 162;

Best Local Similarity 25.0%; Pred. No. 4e+02; Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
Db 107 CATTATTC 114

RESULT 40

ID ADP31269 standard; protein, 165 AA.

AC ADP31269;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2036.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

XX WO2004035732-A2.
PN
XX
PD 29-APR-2004.
PF 28-AUG-2003; 2003WO-US026780.
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406649P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410944P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0410963P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Heestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang WM, Kochakota S, Haisan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as Proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 3267; 428pp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytosstatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
SQ Sequence 165 AA:

Query Match 75.0%; Score 18; DB 8; Length 165;
Best Local Similarity 25.0%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Cy 1 CXXXXXXC 8
Db 62 CTATTTTC 69

RESULT 41
ADP31099
ID ADP31099 standard; protein; 168 AA.
XX
AC ADP31099;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1866.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
FN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406649P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476611P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493379P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Heestir K, Beaurang PA, Behrens D,
PI Halembeck RF, Huang MM, Kothakota S, Halsham L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3097; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 168 AA;
Query Match 75.0%; Score 18; DB 8; Length 168;
Best Local Similarity 25.0%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CXXXXXXC 8
DB 25 CATTTTTC 32
RESULT 42
ADP30794
ID ADP30794 standard; protein; 171 AA.
XX
XX ADP30794;
AC
XX
DT 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #1561.
DE
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
OS
XX
XX WO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PF
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX PA
XX PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halsebeck RF, Huang MM, Kothakota S, Halsehan L, Himmemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX DR WPI; 2004-348438/32.
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX PT genetic, bacterial and viral diseases.
XX PS Claim 1; SEQ ID NO 2792; 428pp; English.
XX CC The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cytostatic,
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPWEB and is not in the specification.
XX SQ Sequence 171 AA;
OY Query March 75 0%; Score 18; DB 8; Length 171;
Best Local Similarity 25.0%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Db 1 CXXXXXXC 8
161 CATATATC 168
RESULT 43
ADP30793
ID ADP30793 standard; protein; 171 AA.
XX AC ADP30793;
XX DT 12-AUG-2004 (first entry)

XX DE Human secreted protein SEQ ID #1560.
XX KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX KW cancer; inflammatory; immune; human secreted protein.
XX OS Homo sapiens.
XX PN WO2004035732-A2.
XX PD 29-APR-2004.
XX PF 28-AUG-2003; 2003WO-US026780.
XX PR 29-AUG-2002; 2002US-0406576P.
XX PR 29-AUG-2002; 2002US-0406579P.
XX PR 29-AUG-2002; 2002US-0406585P.
XX PR 29-AUG-2002; 2002US-0406588P.
XX PR 29-AUG-2002; 2002US-0406608P.
XX PR 29-AUG-2002; 2002US-0406611P.
XX PR 29-AUG-2002; 2002US-0406612P.
XX PR 29-AUG-2002; 2002US-0406616P.
XX PR 29-AUG-2002; 2002US-0406640P.
XX PR 29-AUG-2002; 2002US-0406642P.
XX PR 29-AUG-2002; 2002US-0406646P.
XX PR 29-AUG-2002; 2002US-0406653P.
XX PR 29-AUG-2002; 2002US-0406655P.
XX PR 29-AUG-2002; 2002US-0406666P.
XX PR 17-SEP-2002; 2002US-0410946P.
XX PR 17-SEP-2002; 2002US-0410947P.
XX PR 17-SEP-2002; 2002US-0410949P.
XX PR 17-SEP-2002; 2002US-0410953P.
XX PR 17-SEP-2002; 2002US-0410957P.
XX PR 17-SEP-2002; 2002US-0410958P.
XX PR 17-SEP-2002; 2002US-0410959P.
XX PR 17-SEP-2002; 2002US-0410960P.
XX PR 17-SEP-2002; 2002US-0410961P.
XX PR 17-SEP-2002; 2002US-0410962P.
XX PR 17-SEP-2002; 2002US-0411019P.
XX PR 17-SEP-2002; 2002US-0411022P.
XX PR 17-SEP-2002; 2002US-0411023P.
XX PR 17-SEP-2002; 2002US-0411024P.
XX PR 17-SEP-2002; 2002US-0411032P.
XX PR 17-SEP-2002; 2002US-0411035P.
XX PR 17-SEP-2002; 2002US-0411037P.
XX PR 17-SEP-2002; 2002US-0411041P.
XX PR 17-SEP-2002; 2002US-0411045P.
XX PR 17-SEP-2002; 2002US-0411046P.
XX PR 17-SEP-2002; 2002US-0411052P.
XX PR 17-SEP-2002; 2002US-0411055P.
XX PR 17-SEP-2002; 2002US-0411073P.
XX PR 17-SEP-2002; 2002US-0411082P.
XX PR 17-SEP-2002; 2002US-0411101P.
XX PR 17-SEP-2002; 2002US-0463700P.
XX PR 18-APR-2003; 2003US-0463708P.
XX PR 18-APR-2003; 2003US-0463716P.
XX PR 18-APR-2003; 2003US-0463732P.
XX PR 02-MAY-2003; 2003US-0467199P.
XX PR 02-MAY-2003; 2003US-0467203P.
XX PR 02-MAY-2003; 2003US-0467203P.
XX PR 02-MAY-2003; 2003US-0467230P.
XX PR 19-MAY-2003; 2003US-0471306P.
XX PR 19-MAY-2003; 2003US-0471336P.
XX PR 22-MAY-2003; 2003US-0472420P.
XX PR 09-JUN-2003; 2003US-0476609P.
XX PR 09-JUN-2003; 2003US-0476641P.
XX PR 08-JUL-2003; 2003US-0485218P.
XX PR 08-JUL-2003; 2003US-0485223P.
XX PR 08-JUL-2003; 2003US-0485224P.

PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Hakenbeck RF, Huang MM, Kothakota S, Hsiehan L, Innemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
DR New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
PS Claim 1; SEQ ID NO 2791; 428bp; English.
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cyostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
XX available on WIPWEB and is not in the specification.
SQ Sequence 171 AA;
Query Match 75.0%; Score 18; DB 8; Length 171;
Best Local Similarity 25.0%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Cy 1 CXXXXXXC 8
Db 161 CATTATTC 168
RESULT 44
ID AAG10193 standard; protein; 177 AA.
XX AAG10193;
XX AAG10193;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 8414.
DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS EPI033405-A2.
PN EPI033405-A2.
XX 06-SEP-2000.
PD 25-FEB-2000; 2000EP-00301439.
PF 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-012180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132040P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0135629P.
PR 27-MAY-1999; 99US-0136021P.
PR 28-MAY-1999; 99US-0136392P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139753P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151301P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154033P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155658P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.

PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160757P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161993P.
PR 28-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 75.0%; Score 18; DB 3; Length 177;
Best Local Similarity 25.0%; Pred. No. 4e+02; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 6;

Oy 1 CXXXXXXC 8
Db 87 CSSSSSC 94

RESULT 45
ADP30755
ID ADP30755 standard; protein; 177 AA.

AC ADP30755;
XX 12-AUG-2004 (first entry)
DT XX
XX

DE Human secreted protein SEQ ID #1522.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

OS WO2004035732-A2.

PN XX
XX 29-APR-2004.

PD XX
PF 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.

XX 17-SEP-2002; 2002US-0410946P.

CC	XX	PR	17-SEP--2002;	2002US-0410947P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0410948P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0410949P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0410953P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0410957P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0410958P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0410959P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0410960P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0410961P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0410962P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411019P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411022P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411023P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411024P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411032P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411052P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411053P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411037P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411041P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411045P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411046P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411048P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411052P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411055P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411073P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411082P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411011P.	PR
CC	XX	PR	17-SEP-2002;	2002US-0411111P.	PR
CC	XX	PR	18-APR-2003;	2003US-0463708P.	PR
CC	XX	PR	18-APR-2003;	2003US-0463708P.	PR
CC	XX	PR	18-APR-2003;	2003US-0463716P.	PR
CC	XX	PR	18-APR-2003;	2003US-0463732P.	PR
CC	XX	PR	02-MAY-2003;	2003US-0467199P.	PR
CC	XX	PR	02-MAY-2003;	2003US-0467201P.	PR
CC	XX	PR	02-MAY-2003;	2003US-0467203P.	PR
CC	XX	PR	02-MAY-2003;	2003US-0467230P.	PR
CC	XX	PR	19-MAY-2003;	2003US-0471306P.	PR
CC	XX	PR	19-MAY-2003;	2003US-0471336P.	PR
CC	XX	PR	22-MAY-2003;	2003US-0472420P.	PR
CC	XX	PR	22-MAY-2003;	2003US-0472430P.	PR
CC	XX	PR	09-JUN-2003;	2003US-0476609P.	PR
CC	XX	PR	09-JUN-2003;	2003US-0476641P.	PR
CC	XX	PR	08-JUL-2003;	2003US-0485218P.	PR
CC	XX	PR	08-JUL-2003;	2003US-0485223P.	PR
CC	XX	PR	08-JUL-2003;	2003US-0485224P.	PR
CC	XX	PR	08-JUL-2003;	2003US-0485355P.	PR
CC	XX	PR	14-JUL-2003;	2003US-0486446P.	PR
CC	XX	PR	14-JUL-2003;	2003US-0486480P.	PR
CC	XX	PR	15-JUL-2003;	2003US-0486891P.	PR
CC	XX	PR	15-JUL-2003;	2003US-0486960P.	PR
CC	XX	PR	08-AUG-2003;	2003US-0493341P.	PR
CC	XX	PR	08-AUG-2003;	2003US-0493370P.	PR
CC	XX	PR	08-AUG-2003;	2003US-0493572P.	PR
CC	XX	PR	08-AUG-2003;	2003US-0493577P.	PR
CC	XX	PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.		PA
CC	XX	PI	Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D,		PI
CC	XX	PI	Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T,		PI
CC	XX	PI	Plecece K, Wang Y, Wong JGP, Wu G, Zhang H;		PI
CC	XX	DR	WPI; 2004-348438/32.		DR
CC	XX	PT	New nucleic acid molecule for diagnosing, preventing or treating diseases		PT
CC	XX	PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,		PT
CC	XX	PT	genetic, bacterial and viral diseases.		PT
CC	XX	PS	Claim 1; SEQ ID NO 2753; 428bp; English.		PS
CC	XX	CC	The present invention relates to an isolated nucleic acid molecule		CC
CC	XX	CC	encoding a polypeptide which is believed to be cytostatic,		CC
CC	XX	CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The		CC
CC	XX	CC	composition and methods are useful for diagnosing, preventing and		CC
CC	XX	CC	treating diseases such as proliferative (e.g. cancer), inflammatory,		CC

CC	immune, metabolic, genetic, bacterial and viral diseases. The present sequence represents a human secreted protein. The present sequence is available on WIPoWEB and is not in the specification.
CC	
CC	
XX	
SQ	Sequence 177 AA;
OY	1 CXXXXXXX 8
DB	145 CTATATTTC 152
RESULT 46	
ADP31682	
ID	ADP31682 standard; protein, 177 AA.
XX	
AC	ADP31682;
DT	
XX	12-AUG-2004 (first entry)
DE	Human secreted protein SEQ ID #2449.
XX	
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
RN	cancer; inflammatory; immune; human secreted protein.
OS	Homo sapiens.
XX	
PM	WO2004035732-A2.
PD	
XX	29-APR-2004.
PF	28-AUG-2003; 2003WO-US026780.
PR	29-AUG-2002; 2002US-0406576P.
PR	29-AUG-2002; 2002US-0406579P.
PR	29-AUG-2002; 2002US-0406585P.
PR	29-AUG-2002; 2002US-0406588P.
PR	29-AUG-2002; 2002US-0406608P.
PR	29-AUG-2002; 2002US-0406611P.
PR	29-AUG-2002; 2002US-0406612P.
PR	29-AUG-2002; 2002US-0406616P.
PR	29-AUG-2002; 2002US-0406640P.
PR	29-AUG-2002; 2002US-0406642P.
PR	29-AUG-2002; 2002US-0406646P.
PR	29-AUG-2002; 2002US-0406653P.
PR	29-AUG-2002; 2002US-0406655P.
PR	29-AUG-2002; 2002US-0406666P.
PR	17-SEP-2002; 2002US-0410946P.
PR	17-SEP-2002; 2002US-0410947P.
PR	17-SEP-2002; 2002US-0410948P.
PR	17-SEP-2002; 2002US-0410949P.
PR	17-SEP-2002; 2002US-0410953P.
PR	17-SEP-2002; 2002US-0410957P.
PR	17-SEP-2002; 2002US-0410958P.
PR	17-SEP-2002; 2002US-0410959P.
PR	17-SEP-2002; 2002US-0410960P.
PR	17-SEP-2002; 2002US-0410961P.
PR	17-SEP-2002; 2002US-0410962P.
PR	17-SEP-2002; 2002US-0411019P.
PR	17-SEP-2002; 2002US-0411022P.
PR	17-SEP-2002; 2002US-0411023P.
PR	17-SEP-2002; 2002US-0411024P.
PR	17-SEP-2002; 2002US-0411032P.
PR	17-SEP-2002; 2002US-0411035P.
PR	17-SEP-2002; 2002US-0411037P.
PR	17-SEP-2002; 2002US-0411041P.
PR	17-SEP-2002; 2002US-0411045P.
PR	17-SEP-2002; 2002US-0411046P.
PR	17-SEP-2002; 2002US-0411048P.
PR	17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-041105SP.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476009P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Halshan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3680; 428bp; English.
XX
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WtOWB and is not in the specification.
XX
XX Sequence 177 AA:

SQ

Query Match 75.0%; Score 18; DB 8; Length 177;
Best Local Similarity 25.0%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
Db 39 CAAAAATC 46

RESULT 47

ADT56379

ID ADT56379 standard; protein; 177 AA.

XX AC ADT56379;

DT 13-JAN-2005 (first entry)

XX
XX Plant polypeptide, SEQ ID 6456.

DE
XX plant; transgenic; cold tolerance; growth rate; drought tolerance;
KW disease resistance; galactomannan production; plant growth regulator;
KW heat tolerance; herbicide tolerance; lignin production;
KW extreme osmotic condition tolerance; pathogens resistance;
KW pest resistance; yield improvement; seed oil yield; seed protein yield.

XX

OS Virdidplantae.

XX

FN US2004216190-A1.

XX

PD 28-OCT-2004.

XX

PP 18-DEC-2003; 2003US-00739930.

XX

PR 28-APR-2003; 2003US-00424599.

XX

PR 28-APR-2003; 2003US-00425115.

XX

PA (KOVA/) KOVALIC D K.

XX

PI Kovalic DK;

XX

DR WPI; 2004-757369/74.

XX

XX

PT New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
XX biological characteristics.

XX

PS Claim 2; SEQ ID NO 6456; 14bp; English.

XX

XX

CC The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (CDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The CDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomannan production, for production
CC of plant growth regulator, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant protein
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPRO at
CC seqdata.uspro.gov/sequence.html?DocID=20040216190.

XX

SQ Sequence 177 AA;

Query Match 75.0%; Score 18; DB 8; Length 177;

Best Local Similarity 25.0%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
DB 87 CXXXXXXC 94

RESULT 48

ABM86593
ID ABM86593 standard; protein, 182 AA.

XX AC ABM86593;

XX DT 02-JUN-2005 (first entry)

DE Rice abiotic stress responsive polypeptide SEQ ID NO:4839.

KM abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.

XX OS Oryza sativa.

XX PN MO2003008540-A2.

XX PD 30-JAN-2003.

XX PF 21-JUN-2002; 2002WO-US019668.

XX PR 22-JUN-2001; 2001US-0300112P.

PR 24-AUG-2001; 2001US-0314662P.

PR 26-SEP-2001; 2001US-0325277P.

PR 21-NOV-2001; 2001US-0332132P.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Krepis J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;

XX PI Moughamer T, Provart N, Ricke D, Zhu T;

XX DR WPI; 2003-248011/24.

XX PT New stress-responsive nucleic acid, useful for altering the

XX PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold

XX PT stress, salt stress or osmotic stress.

XX PS Claim 1; SEQ ID NO 4839; 89pp; English.

XX CC The invention relates to novel abiotic stress responsive polynucleotides

XX CC and polypeptides. Also disclosed are vectors, expression cassettes, host

XX CC cells, and plants containing such polynucleotides. Also disclosed are

XX CC methods for using the polynucleotides and polypeptides to alter the

XX CC responsiveness of a plant to abiotic stress. The invention is useful in

XX CC agriculture. The nucleic acid is useful for determining whether a test

XX CC plant has been exposed to an abiotic stress condition. It is also useful

XX CC for selecting an agent that alters abiotic stress regulated

XX CC polynucleotide expression in a plant cell, and to identify a homolog or

XX CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid

XX CC molecule and the polypeptide encoded by it are useful in altering the

XX CC responsiveness of a plant to an abiotic stress, such as cold stress, salt

XX CC stress, osmotic stress or any of their combinations. The present sequence

XX CC is used in the exemplification of the invention

XX SQ Sequence 182 AA;

Query Match 75.0%; Score 18; DB 7; Length 182;

Best Local Similarity 25.0%; Pred. No. 4e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8

DB 15 CSTASAC 22

RESULT 49

ABM89559
ID ABM89559 standard; protein, 185 AA.

XX AC ABM89559;

XX DT 02-JUN-2005 (first entry)

DE Rice abiotic stress responsive polypeptide SEQ ID NO:7805.

KM abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.

XX OS Oryza sativa.

XX PN MO2003008540-A2.

XX PD 30-JAN-2003.

XX PF 21-JUN-2002; 2002WO-US019668.

XX PR 22-JUN-2001; 2001US-0300112P.

PR 24-AUG-2001; 2001US-0314662P.

PR 26-SEP-2001; 2001US-0325277P.

PR 21-NOV-2001; 2001US-0332132P.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Krepis J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;

XX PI Moughamer T, Provart N, Ricke D, Zhu T;

XX DR WPI; 2003-248011/24.

XX PT New stress-responsive nucleic acid, useful for altering the

XX PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold

XX PT stress, salt stress or osmotic stress.

XX PS Claim 1; SEQ ID NO 7805; 89pp; English.

XX CC The invention relates to novel abiotic stress responsive polynucleotides

XX CC and polypeptides. Also disclosed are vectors, expression cassettes, host

XX CC cells, and plants containing such polynucleotides. Also disclosed are

XX CC methods for using the polynucleotides and polypeptides to alter the

XX CC responsiveness of a plant to abiotic stress. The invention is useful in

XX CC agriculture. The nucleic acid is useful for determining whether a test

XX CC plant has been exposed to an abiotic stress condition. It is also useful

XX CC for selecting an agent that alters abiotic stress regulated

XX CC polynucleotide expression in a plant cell, and to identify a homolog or

XX CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid

XX CC molecule and the polypeptide encoded by it are useful in altering the

XX CC responsiveness of a plant to an abiotic stress, such as cold stress, salt

XX CC stress, osmotic stress or any of their combinations. The present sequence

XX CC is used in the exemplification of the invention

XX SQ Sequence 185 AA;

Query Match 75.0%; Score 18; DB 7; Length 185;

Best Local Similarity 25.0%; Pred. No. 4.1e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8

DB 148 CSTATTSC 155

RESULT 50

ADP31576
ID ADP31576 standard; protein, 192 AA.

XX AC ADP31576;

XX DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2343.

KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX Homo sapiens.
XX MO2004035732-A2.
XX
PD 29-APR-2004.
XX
PE 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476099P.
PR 09-JUN-2003; 2003US-0476411P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,
XX
DR WPI; 2004-348438/32.

XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

XX
PS Claim 1; SEQ ID NO 3574; 428bp; English.

XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytosratic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.

XX
SQ Sequence 192 AA;

Query Match 75.0%; Score 18; DB 8; Length 192;
Best Local Similarity 25.0%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
DB 97 CTAAATC 104

Search completed: January 4, 2006, 15:56:14
Job time : 56.7478 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 ; Search time: 8.27626 seconds
(without alignments)
92.983 Million cell updates/sec

Title: US-09-932-322-9
Perfect score: 24
Sequence: 1 CXXXXXXC 8

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 1000 summaries

Database :
1: PIR_80:*
2: pir1:*
3: pir2:*
4: pir3:*
5: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	18	75.0	177 2 P96719	unknown protein, 3
2	18	75.0	355 2 A34597	rhamboind protein -
3	18	75.0	367 2 C39590	TPA-induced protei
4	18	75.0	492 2 S49147	ERF-2 protein - hu
5	18	75.0	843 1 JDUVLJ1	DNA-directed DNA p
6	18	75.0	843 1 JDUVLJ2	DNA-directed DNA p
7	18	75.0	1743 2 T26859	hypothetical prote
8	18	75.0	2543 2 T31687	hypothetical prote
9	18	75.0	26 2 A44036	surface antigen - p
10	17	70.8	89 2 S72742	collagen alpha 1(X
11	17	70.8	101 2 T42262	Bl177 F1 32 protei
12	17	70.8	120 2 S72979	hypothetical prote
13	17	70.8	122 2 J00150	probable L-a-amino
14	17	70.8	123 2 S25574	hypothetical 13k p
15	17	70.8	135 1 RWHUVY	Ig heavy chain V r
16	17	70.8	139 2 A71123	T-cell receptor be
17	17	70.8	145 2 H87436	hypothetical prote
18	17	70.8	182 2 T22305	hypothetical prote
19	17	70.8	203 2 E70641	hypothetical prote
20	17	70.8	212 2 H75277	hypothetical prote
21	17	70.8	224 2 A41128	myogenin - human
22	17	70.8	230 2 S09778	hypothetical prote
23	17	70.8	237 2 P95411	hypothetical prote
24	17	70.8	238 1 JH0367	ribonuclease (EC 3
25	17	70.8	243 2 A35871	trypsin (EC 3.4.21
26	17	70.8	248 2 T47902	hypothetical prote
27	17	70.8	267 2 E87269	exopolysaccharide
28	17	70.8	271 2 F70771	probable glutamate
29	17	70.8	272 1 S72790	glutamate racemase

30	17	70.8	284 2	I51172	transcription fact
31	17	70.8	292 2	C64795	ctlg protein - Esc
32	17	70.8	292 2	D90710	2-(5')-triphosphor
33	17	70.8	292 2	H85560	2-(5')-triphosphor
34	17	70.8	297 2	A98294	hypothetical prote
35	17	70.8	297 2	AH2989	permease (imported
36	17	70.8	305 2	AE0857	conserved hypotet
37	17	70.8	305 2	G65056	hypothetical prote
38	17	70.8	307 2	H85925	hypothetical prote
39	17	70.8	307 2	A91080	hypothetical prote
40	17	70.8	309 2	T32376	hypothetical prote
41	17	70.8	347 2	B53522	20k cyclophilin -
42	17	70.8	348 2	T36890	probable iron-side
43	17	70.8	350 2	C84548	hypothetical prote
44	17	70.8	352 2	B87507	conserved hypotet
45	17	70.8	365 2	G71287	conserved hypotet
46	17	70.8	376 2	D97175	nifS family enzyme
47	17	70.8	400 2	C34443	nitrogenase cofact
48	17	70.8	400 2	AE1988	nitrogenase cofact
49	17	70.8	408 1	MMNVIA	immediate-early pr
50	17	70.8	408 2	A72869	early gene transac
51	17	70.8	412 2	AG3637	cis,cis-muconate t
52	17	70.8	419 2	C40901	hypothetical prote
53	17	70.8	429 2	T16656	hypothetical prote
54	17	70.8	439 2	A36385	surface antigen se
55	17	70.8	456 2	S20597	ribonuclease inhib
56	17	70.8	461 2	A18588	ribonuclease-angio
57	17	70.8	467 2	JC6150	Gram-negative bact
58	17	70.8	471 2	AS4840	variant surface gl
59	17	70.8	471 2	S18446	variant surface gl
60	17	70.8	489 2	S71955	MDM2-like p53-bind
61	17	70.8	492 1	MMAD55	early E1B 55K prot
62	17	70.8	539 2	H84640	probable Rieseke ir
63	17	70.8	541 2	T15299	hypothetical prote
64	17	70.8	544 2	S44814	F44B9.3 protein -
65	17	70.8	547 1	S51475	cytochrome P450 cy
66	17	70.8	567 2	T49942	hypothetical prote
67	17	70.8	577 2	B37057	integrin beta-6 ch
68	17	70.8	615 2	B86212	hypothetical prote
69	17	70.8	631 2	T00925	hypothetical prote
70	17	70.8	638 2	D86477	protein P1504.27 l
71	17	70.8	655 2	A59430	hypothetical prote
72	17	70.8	658 2	T19487	hypothetical prote
73	17	70.8	662 2	T17211	trophozoite surfac
74	17	70.8	667 2	A48579	white protein - fr
75	17	70.8	687 1	FYEFW	integrin beta-6 ch
76	17	70.8	788 2	A37057	DNA-directed DNA p
77	17	70.8	843 1	JDUVLJ3	probable nirB prot
78	17	70.8	853 2	H70939	complement C6 prec
79	17	70.8	934 1	A34372	hypothetical prote
80	17	70.8	1028 2	T34360	hypothetical prote
81	17	70.8	1156 2	T23308	hypothetical prote
82	17	70.8	1279 2	A47363	RNA helicase A - h
83	17	70.8	1287 2	T46032	nuclear DNA helica
84	17	70.8	1345 2	S55659	legumain protein 7
85	17	70.8	1474 2	D88550	protein ZC84.6 (im
86	17	70.8	1474 2	H88380	protein T22F7.3 (i
87	17	70.8	1522 2	H88380	collagen alpha 1(X
88	17	70.8	1747 2	A45974	collagen alpha 1(X
89	17	70.8	1857 2	S31212	collagen alpha 1(X
90	17	70.8	1888 2	S78476	collagen alpha 1(X
91	17	70.8	2395 1	S50820	surface protein ty
92	17	70.8	2533 2	T28675	alpha-51D immobili
93	17	70.8	2533 2	T28674	alpha-51D immobili
94	17	70.8	2844 2	S28291	hypothetical prote
95	17	70.8	3124 2	A40020	collagen alpha 1(X
96	17	70.8	3229 2	S27852	probable cell-surf
97	16	66.7	38 2	A47307	heat-stable entero
98	16	66.7	39 2	B41933	meting pheromone E
99	16	66.7	44 2	S38275	toxin PLTX-II - sp
100	16	66.7	45 1	PFY2	fulvicin C - Myxoc
101	16	66.7	55 1	NTN82	neurotoxin B-II -
102	16	66.7	57 2	E82522	hypothetical prote
				A46654	growth modulatory

103	16	66.7	57	2	B46654	growth modulatory
104	16	66.7	57	2	C46654	growth modulatory
105	16	66.7	62	1	S54336	metallothionein-2a
106	16	66.7	62	2	S54335	metallothionein-2c
107	16	66.7	72	1	WMBP2P	ogr protein - phag
108	16	66.7	72	2	D82519	hypothetical prote
109	16	66.7	79	2	G82811	hypothetical prote
110	16	66.7	83	2	S11843	viral protein - Ag
111	16	66.7	83	2	AB3251	vira/g regulated p
112	16	66.7	91	2	G90824	hypothetical prote
113	16	66.7	92	2	D37057	epithelial cell gl
114	16	66.7	92	2	T30632	probable vitron co
115	16	66.7	100	2	A13337	hypothetical prote
116	16	66.7	104	2	B81045	hypothetical prote
117	16	66.7	109	2	S24253	Ig heavy chain V r
118	16	66.7	114	2	H71018	hypothetical prote
119	16	66.7	115	2	B96664	probable RING zinc
120	16	66.7	118	1	BSSNKL	phospholipase A2 (
121	16	66.7	118	1	BSSNKL	phospholipase A2 (
122	16	66.7	118	2	D34860	phospholipase A2 (
123	16	66.7	118	2	E34860	phospholipase A2 (
124	16	66.7	118	2	G34860	phospholipase A2 (
125	16	66.7	118	2	P34860	phospholipase A2 (
126	16	66.7	118	2	H34860	phospholipase A2 (
127	16	66.7	119	1	PSNOA1	phospholipase A2 h
128	16	66.7	119	2	G97003	hypothetical prote
129	16	66.7	120	2	E70424	plin - Aquifex ae
130	16	66.7	125	2	JH0184	hypothetical prote
131	16	66.7	126	2	D72784	hypothetical prote
132	16	66.7	132	2	I39193	gene HOXA1 protein
133	16	66.7	133	2	B30242	stem cell protein
134	16	66.7	135	2	D90847	hypothetical prote
135	16	66.7	136	2	I50554	gonadotropin II be
136	16	66.7	136	2	E71883	hypothetical prote
137	16	66.7	138	2	A80258	conserved hypotnet
138	16	66.7	140	2	T49481	hypothetical prote
139	16	66.7	140	2	T33359	hypothetical prote
140	16	66.7	141	2	T49380	hypothetical prote
141	16	66.7	141	2	PC1294	trichozoite surfac
142	16	66.7	142	2	J50510	fusaric acid resis
143	16	66.7	143	1	QOEC32	hypothetical prote
144	16	66.7	144	1	PSHUPP	phospholipase A2 (
145	16	66.7	144	2	UN0480	phospholipase A2 (
146	16	66.7	144	2	S01391	phospholipase A2 (
147	16	66.7	145	1	A60480	interleukin-9 prec
148	16	66.7	145	1	PSKFXU	phospholipase A2 (
149	16	66.7	145	2	S07983	phospholipase A2 h
150	16	66.7	145	2	S01390	phospholipase A2 (
151	16	66.7	145	2	S74292	hypothetical prote
152	16	66.7	145	2	T42280	hypothetical prote
153	16	66.7	145	2	T15608	hypothetical prote
154	16	66.7	148	1	PSHU	phospholipase A2 (
155	16	66.7	149	2	C81959	probable pilin NMA
156	16	66.7	150	2	C42744	exonuclease (EC 3.
157	16	66.7	151	2	T25047	hypothetical prote
158	16	66.7	153	2	S34255	hypothetical prote
159	16	66.7	157	2	S55956	probable membrane
160	16	66.7	160	2	J00542	185K secretory pro
161	16	66.7	166	2	P84449	hypothetical prote
162	16	66.7	172	2	AB3514	outer membrane pro
163	16	66.7	173	2	T14827	lipid transfer pro
164	16	66.7	174	2	AC0306	probable outer mem
165	16	66.7	177	2	G86917	conserved hypotnet
166	16	66.7	177	2	D70654	hypothetical prote
167	16	66.7	178	2	G71195	hypothetical prote
168	16	66.7	181	2	UN0342	N-methyl-D-asparta
169	16	66.7	181	2	T15412	knutz type subcil
170	16	66.7	183	2	JX0311	knutz type subcil
171	16	66.7	183	2	I38923	microfibril-associ
172	16	66.7	183	2	A54151	microfibril-associ
173	16	66.7	183	2	A49313	microfibril-associ
174	16	66.7	183	2	S55599	hypothetical prote
175	16	66.7	186	2	B82625	outer membrane pro
176	16	66.7	187	2	S36918	zygote-specific pr
177	16	66.7	187	2	P90315	hypothetical prote
178	16	66.7	188	2	T33886	hypothetical prote
179	16	66.7	189	2	G70347	phenylacrylic acid
180	16	66.7	190	2	JX0310	knutz type subcil
181	16	66.7	190	2	A59416	knutz type subcil
182	16	66.7	190	2	A59418	knutz type subcil
183	16	66.7	195	2	B82857	PilX protein XPO03
184	16	66.7	198	2	S83435	conserved hypotnet
185	16	66.7	199	2	S35441	serine-tRNA ligase
186	16	66.7	201	2	S23324	gene achate prote
187	16	66.7	202	2	AC2875	hypothetical prote
188	16	66.7	204	2	T28939	hypothetical prote
189	16	66.7	205	2	T37947	hypothetical prote
190	16	66.7	207	2	B81418	hypothetical prote
191	16	66.7	208	2	S28675	hypothetical prote
192	16	66.7	208	2	P00278	hypothetical prote
193	16	66.7	208	2	B64925	probable oxidoredu
194	16	66.7	208	2	B90926	probable oxidoredu
195	16	66.7	208	2	A85775	urease accessory p
196	16	66.7	213	1	S71282	ribosomal protein
197	16	66.7	220	1	S71282	hypothetical prote
198	16	66.7	220	2	T21730	hypothetical prote
199	16	66.7	222	2	I39192	gene HOXA1 protein
200	16	66.7	223	2	A65172	hypothetical prote
201	16	66.7	224	2	A36675	type 4 prepilin-11
202	16	66.7	225	2	B65127	hypothetical prote
203	16	66.7	225	2	T49769	hypothetical prote
204	16	66.7	225	2	T33361	myogenin - chicken
205	16	66.7	227	2	B35882	myogenin - chicken
206	16	66.7	229	1	TRBOTR	hypothetical prote
207	16	66.7	229	1	T29203	hypothetical prote
208	16	66.7	231	1	TRPGRR	trypsin (EC 3.4.21
209	16	66.7	231	2	T02765	trypsin (EC 3.4.21
210	16	66.7	233	2	S71464	trypsin (EC 3.4.21
211	16	66.7	234	2	T52018	conserved hypotnet
212	16	66.7	237	2	S08073	hypothetical prote
213	16	66.7	237	2	T25152	gene 2 protein - h
214	16	66.7	238	1	W2B82	trypsin (EC 3.4.21
215	16	66.7	238	2	S31779	trypsin (EC 3.4.21
216	16	66.7	240	2	S39047	trypsin (EC 3.4.21
217	16	66.7	240	2	H82289	conserved hypotnet
218	16	66.7	240	2	T25641	hypothetical prote
219	16	66.7	240	2	A43912	myogenin - Japanes
220	16	66.7	241	2	S39048	trypsin (EC 3.4.21
221	16	66.7	242	2	C70895	hypothetical prote
222	16	66.7	242	2	T45795	hypothetical prote
223	16	66.7	245	2	T24565	hypothetical prote
224	16	66.7	247	2	S13813	trypsin (EC 3.4.21
225	16	66.7	247	2	S05494	trypsin (EC 3.4.21
226	16	66.7	248	2	S55067	trypsin (EC 3.4.21
227	16	66.7	248	2	T21786	hypothetical prote
228	16	66.7	248	2	F85067	hypothetical prote
229	16	66.7	249	2	T32060	hypothetical prote
230	16	66.7	254	2	G71545	probable deoxyoctu
231	16	66.7	254	2	A56447	CMP-2-Xeto-3-deoxy
232	16	66.7	258	2	T32542	hypothetical prote
233	16	66.7	259	2	T46881	succinate dehydrog
234	16	66.7	261	2	F82728	succinate dehydrog
235	16	66.7	262	2	F72003	probable outer mem
236	16	66.7	262	2	G86619	probable omp limpo
237	16	66.7	262	2	F87253	dieneolactone hydro
238	16	66.7	262	2	A57391	TSL-2 protein pre
239	16	66.7	265	2	A95401	protein limported
240	16	66.7	265	2	S06462	beta-lactamase (EC
241	16	66.7	266	2	S33163	beta-lactamase (EC
242	16	66.7	267	2	S22684	beta-lactamase (EC
243	16	66.7	268	2	T10304	inhibitor of apopt
244	16	66.7	268	2	A53989	apoptosis-inhibiti
245	16	66.7	272	1	A36082	insulin-like growt
246	16	66.7	274	2	T47933	hypothetical prote
247	16	66.7	276	2	T09131	chitinase (EC 3.2.
248	16	66.7	279	2	C70458	diaminopimelate ep

249	16	66.7	280	2	T09504	322	16	66.7	366	2	A95382	probable iron upta
250	16	66.7	283	2	A64179	323	16	66.7	367	2	T49048	hypothetical prote
251	16	66.7	287	2	T15774	324	16	66.7	369	2	AE2345	phospho-N-acetylmu
252	16	66.7	288	2	T04401	325	16	66.7	369	2	G81101	membrane-bound lyt
253	16	66.7	288	2	T21790	326	16	66.7	369	2	T15213	hydrogenase expres
254	16	66.7	288	2	T38075	327	16	66.7	370	2	B64632	hydrogenase expres
255	16	66.7	289	1	Q2AD2	328	16	66.7	370	2	F71883	hydrogenase expres
256	16	66.7	289	1	Q2AD5	329	16	66.7	371	2	A39234	opsin - bluebotle
257	16	66.7	292	2	T23966	330	16	66.7	372	2	C97791	nfs protein homol
258	16	66.7	293	2	T03883	331	16	66.7	373	1	COOF	riodopsin - fruit
259	16	66.7	294	2	T23682	332	16	66.7	374	2	D83328	probable aminotran
260	16	66.7	296	2	S77427	333	16	66.7	374	2	S40691	opsin rh1 - fruit
261	16	66.7	297	2	A10827	334	16	66.7	374	2	T05660	hypothetical prote
262	16	66.7	297	2	T38161	335	16	66.7	374	2	T18442	zinc finger protei
263	16	66.7	298	2	AF0578	336	16	66.7	376	2	B16522	iron-sulfur cofact
264	16	66.7	302	2	B71310	337	16	66.7	377	2	C89456	protein F55A4.8 [i
265	16	66.7	303	2	T19289	338	16	66.7	378	2	C89456	protein F55A4.8 [i
266	16	66.7	304	2	B83052	339	16	66.7	378	2	A82406	Gung protein d ubiq
267	16	66.7	305	2	T49720	340	16	66.7	379	2	S67856	hypothetical prote
268	16	66.7	307	1	GSFF3	341	16	66.7	379	2	F71553	conserved hypothet
269	16	66.7	309	2	S41427	342	16	66.7	381	2	A35300	G protein-coupled
270	16	66.7	309	2	S37028	343	16	66.7	383	2	I53870	Bdg-1 orphan recep
271	16	66.7	309	2	S39957	344	16	66.7	384	2	S68410	lipase Arab-1 - Ar
272	16	66.7	309	2	C95977	345	16	66.7	384	2	F97711	hypothetical prote
273	16	66.7	311	2	S20085	346	16	66.7	384	2	T49084	hypothetical prote
274	16	66.7	312	1	QOBE07	347	16	66.7	385	2	T04725	hypothetical prote
275	16	66.7	313	2	D85855	348	16	66.7	385	2	T06484	aspartate carboxy
276	16	66.7	313	2	B91011	349	16	66.7	387	2	C41880	nitrogenase cofact
277	16	66.7	313	2	B49885	350	16	66.7	388	2	T16861	hypothetical prote
278	16	66.7	314	2	T27686	351	16	66.7	389	2	T23167	hypothetical prote
279	16	66.7	314	2	B43652	352	16	66.7	391	1	J01626	actachment protein
280	16	66.7	318	2	B91177	353	16	66.7	391	2	T08338	conserved hypothet
281	16	66.7	319	2	JC2252	354	16	66.7	392	2	B47071	nitrogenase cofact
282	16	66.7	321	2	T02236	355	16	66.7	393	2	F82657	beta-ketocacyl - [ACP
283	16	66.7	322	2	S54806	356	16	66.7	399	2	T21015	hypothetical prote
284	16	66.7	323	1	A48149	357	16	66.7	404	2	A46274	HIV gp120-binding
285	16	66.7	323	2	T03614	358	16	66.7	411	2	E97338	3-oxocacyl - (acyl - ca
286	16	66.7	323	2	S56777	359	16	66.7	411	2	S58105	Cu metalloregulato
287	16	66.7	328	2	C72370	360	16	66.7	416	2	S27198	homeoic protein H
288	16	66.7	328	2	T36494	361	16	66.7	418	2	T15142	hypothetical prote
289	16	66.7	329	2	UC7935	362	16	66.7	424	2	JC5891	omega 6 desaturase
290	16	66.7	329	2	AB1327	363	16	66.7	425	2	B90113	seryl-tRNA synthet
291	16	66.7	329	2	AD1698	364	16	66.7	425	2	JC5909	AS3 protein - fru
292	16	66.7	330	1	UN0561	365	16	66.7	425	2	A37912	thrombin receptor
293	16	66.7	330	2	I56100	366	16	66.7	427	2	G71272	iron-sulfur cofact
294	16	66.7	330	2	C68023	367	16	66.7	429	2	T21113	hypothetical prote
295	16	66.7	330	2	T19644	368	16	66.7	431	2	T36682	probable secreted
296	16	66.7	331	2	T26004	369	16	66.7	432	2	C70381	large subunit of i
297	16	66.7	331	2	A30242	370	16	66.7	432	2	A43448	thrombin receptor
298	16	66.7	332	2	AD3535	371	16	66.7	435	2	T15143	hypothetical prote
299	16	66.7	334	2	H69214	372	16	66.7	439	2	T28841	hypothetical prote
300	16	66.7	335	2	B71361	373	16	66.7	442	2	T24196	hypothetical prote
301	16	66.7	335	2	G01448	374	16	66.7	442	2	D84600	probable xyulose
302	16	66.7	337	2	A53041	375	16	66.7	444	2	F86329	hypothetical prote
303	16	66.7	338	2	G68986	376	16	66.7	448	2	S41725	integrase - Saccha
304	16	66.7	338	2	T06603	377	16	66.7	449	2	T26571	hypothetical prote
305	16	66.7	340	2	T48288	378	16	66.7	450	2	T17234	hypothetical prote
306	16	66.7	344	1	RWRTC2	379	16	66.7	450	2	AD3117	polylacturonase
307	16	66.7	344	2	G82428	380	16	66.7	453	2	A98170	hypothetical prote
308	16	66.7	344	2	S72723	381	16	66.7	456	2	A31857	ribonuclease inhib
309	16	66.7	345	2	B85881	382	16	66.7	456	2	T37575	bacterial cell cyc
310	16	66.7	345	2	B65012	383	16	66.7	466	2	AD3477	malate dehydrogena
311	16	66.7	345	2	H91036	384	16	66.7	470	2	A40697	63K sperm flagella
312	16	66.7	345	2	A05279	385	16	66.7	474	2	S31712	beta-1,3-glucanase
313	16	66.7	347	2	S41638	386	16	66.7	475	2	S53389	glutamate decarbox
314	16	66.7	348	2	JQ0431	387	16	66.7	475	2	T50686	peroxisomal Ca-dep
315	16	66.7	349	2	S53864	388	16	66.7	477	1	I38893	transcription fact
316	16	66.7	349	2	T16882	389	16	66.7	478	2	S31906	beta-1,3-glucanase
317	16	66.7	355	2	C96651	390	16	66.7	481	2	T20309	hypothetical prote
318	16	66.7	358	2	T00578	391	16	66.7	486	2	JC7212	stretch-activated
319	16	66.7	359	2	D89007	392	16	66.7	486	2	T38087	hypothetical prote
320	16	66.7	359	2	T26813	393	16	66.7	487	2	F84727	hypothetical prote
321	16	66.7	364	2	B36313	394	16	66.7	488	1	JC2385	protein disulfide-

395	16	66.7	492	2	JCS169	alkaline nuclease
396	16	66.7	496	2	D83614	conserved hypobeth
397	16	66.7	498	2	AP0157	probable aldehyde
398	16	66.7	498	2	A48203	interleukin-14 pre
399	16	66.7	499	2	B81914	probable periplasm
400	16	66.7	504	1	A28807	protein disulfide-
401	16	66.7	504	1	S41661	protein disulfide-
402	16	66.7	504	2	G02474	interferon regulat
403	16	66.7	505	1	B69417	probable thymidine
404	16	66.7	505	1	SS5507	protein disulfide-
405	16	66.7	505	1	SS3994	protein disulfide-
406	16	66.7	505	2	SS8363	protein disulfide-
407	16	66.7	505	2	JCS704	hydroxymethylgluta
408	16	66.7	509	2	S46314	probable membrane
409	16	66.7	509	2	H70597	carboxypeptidase-1
410	16	66.7	510	2	T48977	Rtn-1-like protein
411	16	66.7	510	2	H90787	probable zinc meta
412	16	66.7	512	2	T37819	serine-trna ligase
413	16	66.7	514	2	G01026	dopamine receptor-
414	16	66.7	514	2	D56849	hypothetical prote
415	16	66.7	519	2	T45764	phosphatidylinosit
416	16	66.7	521	2	A09100	hypothetical prote
417	16	66.7	521	2	T27192	probable interleuk
418	16	66.7	522	2	S71821	hypothetical prote
419	16	66.7	526	2	T13687	hypothetical prote
420	16	66.7	528	2	T33599	hypothetical prote
421	16	66.7	533	2	H71492	probable hsp-60 -
422	16	66.7	534	2	B81654	60 kDa chaperonin
423	16	66.7	536	2	H85647	hypothetical prote
424	16	66.7	547	1	A32244	60K cysteine-rich
425	16	66.7	547	2	S19607	alkaline phosphata
426	16	66.7	547	2	B43584	60K cysteine-rich
427	16	66.7	548	2	B84306	hypothetical prote
428	16	66.7	553	2	D71515	60K cysteine-rich
429	16	66.7	554	2	C81671	60 kDa outer membr
430	16	66.7	554	2	T25288	hypothetical prote
431	16	66.7	556	1	S12602	60K cysteine-rich
432	16	66.7	556	2	A86560	60 kDa Cysteine-ri
433	16	66.7	557	1	B39439	60K cysteine-rich
434	16	66.7	557	2	D84496	probable retroelem
435	16	66.7	557	2	A48434	variant-specific s
436	16	66.7	558	2	JCS204	60K cysteine-rich
437	16	66.7	561	2	S63388	probable membrane
438	16	66.7	569	2	D85354	hypothetical prote
439	16	66.7	573	2	T29617	related to rna bin
440	16	66.7	577	2	T28807	hypothetical prote
441	16	66.7	579	2	T24880	hypothetical prote
442	16	66.7	589	2	T43210	fibulin-ID precurs
443	16	66.7	591	2	T39195	probable amino aci
444	16	66.7	600	2	B46642	DNA-directed DNA p
445	16	66.7	601	2	B96744	unknown protein [i
446	16	66.7	604	2	T49577	hypothetical prote
447	16	66.7	617	2	T49444	luciferin A related
448	16	66.7	631	2	S71508	dark-type molecula
449	16	66.7	638	2	G02068	white homolog - hu
450	16	66.7	643	2	T25473	hypothetical prote
451	16	66.7	655	2	T51792	hypothetical prote
452	16	66.7	655	2	JCS2005	integrin beta-5 ch
453	16	66.7	657	2	B72486	hypothetical prote
454	16	66.7	668	2	A42275	chemoreceptor prot
455	16	66.7	672	2	T12524	hypothetical prote
456	16	66.7	677	2	C42125	tropocholel cystei
457	16	66.7	680	2	PN0510	integrin beta-3 ch
458	16	66.7	686	2	S30075	ferric reductase (
459	16	66.7	686	2	S43562	KOBE5.3 proteain -
460	16	66.7	689	2	T42760	fibulin, splice fo
461	16	66.7	692	2	T32980	BRCore-NS-Z3 prote
462	16	66.7	704	2	S21911	tran protein homol
463	16	66.7	704	2	T31227	probable Tonb-depe
464	16	66.7	706	2	A81848	hypothetical prote
465	16	66.7	707	2	G86894	hypothetical prote
466	16	66.7	709	2	T29692	hypothetical prote
467	16	66.7	712	2	T42990	fibulin 1, splice
468	16	66.7	713	2	I50128	fibroblast growth
469	16	66.7	713	2	A35502	major surface-labe
470	16	66.7	722	2	B71403	hypothetical prote
471	16	66.7	723	2	PN0509	integrin beta-3 ch
472	16	66.7	734	2	UC4861	feritilin beta cha
473	16	66.7	735	2	G02937	feritilin beta - cr
474	16	66.7	743	2	T02147	hypothetical prote
475	16	66.7	746	2	S62365	SNF1-related prote
476	16	66.7	752	2	T48574	hypothetical prote
477	16	66.7	752	2	T20871	hypothetical prote
478	16	66.7	753	2	B36268	platelet glycoprot
479	16	66.7	758	2	T48815	mixed-linked glucu
480	16	66.7	761	2	A46193	88K B-26-specific
481	16	66.7	763	2	E96693	probable terpene s
482	16	66.7	767	2	T07693	hypothetical prote
483	16	66.7	769	2	SS0966	probable membrane
484	16	66.7	770	2	T40027	nitrile reductase
485	16	66.7	773	2	I46059	beta-1 integrin su
486	16	66.7	778	2	A60798	platelet glycoprot
487	16	66.7	781	2	S43534	integrin beta3 - c
488	16	66.7	788	2	A26547	platelet glycoprot
489	16	66.7	788	2	T77349	platelet glycoprot
490	16	66.7	788	2	I51530	integrin beta-3 su
491	16	66.7	798	2	B27079	fibronectin recept
492	16	66.7	798	2	A28193	integrin beta-1 ch
493	16	66.7	798	2	B28193	integrin beta-1* c
494	16	66.7	798	2	T22793	hypothetical prote
495	16	66.7	799	2	A38308	integrin beta-5 ch
496	16	66.7	801	2	G82302	probable cellulob
497	16	66.7	805	2	I40029	nitrile reductase
498	16	66.7	807	2	B83726	assimilatory nitri
499	16	66.7	810	2	B30848	plasma (EC 3.4.21
500	16	66.7	817	2	T25674	hypothetical prote
501	16	66.7	818	2	JC4058	fibroblast growth
502	16	66.7	822	2	S19947	fibroblast growth
503	16	66.7	822	2	B49151	fibroblast growth
504	16	66.7	822	2	T01622	probable salt-indu
505	16	66.7	823	2	D86165	protein F15K9.3 [i
506	16	66.7	829	2	T45883	fibroblast growth
507	16	66.7	834	2	T09369	hypothetical prote
508	16	66.7	838	2	T04449	hypothetical prote
509	16	66.7	849	2	E86306	similar to tuftell
510	16	66.7	849	2	T01286	probable RNA-bind
511	16	66.7	863	1	S51789	VDL receptor prec
512	16	66.7	872	2	S33015	hypothetical prote
513	16	66.7	884	2	E75489	conserved hypobeth
514	16	66.7	885	2	JN0339	N-methyl-D-asparta
515	16	66.7	889	2	D71414	hypothetical prote
516	16	66.7	897	2	S67283	hypothetical prote
517	16	66.7	901	2	JN0337	N-methyl-D-asparta
518	16	66.7	905	2	T23229	hypothetical prote
519	16	66.7	905	2	JN0341	N-methyl-D-asparta
520	16	66.7	906	2	A46286	N-methyl-D-asparta
521	16	66.7	907	2	B96692	probable wall-abso
522	16	66.7	913	2	G64110	hypothetical prote
523	16	66.7	922	2	JN0340	N-methyl-D-asparta
524	16	66.7	922	2	JN0338	N-methyl-D-asparta
525	16	66.7	935	2	S64384	probable membrane
526	16	66.7	938	2	A46612	N-methyl-D-asparta
527	16	66.7	938	2	S21104	N-methyl-D-asparta
528	16	66.7	938	2	S19710	N-methyl-D-asparta
529	16	66.7	943	2	A47551	N-methyl-D-asparta
530	16	66.7	946	2	T01460	hypothetical prote
531	16	66.7	959	2	JN0336	N-methyl-D-asparta
532	16	66.7	964	2	S48404	probable membrane
533	16	66.7	991	1	P1BVCV	RNA 1 protein - cu
534	16	66.7	992	2	A31666	hypothetical prote
535	16	66.7	1007	2	T01437	hypothetical prote
536	16	66.7	1025	2	B54718	dihydropyrimidine
537	16	66.7	1025	2	A54718	probable membrane
538	16	66.7	1029	2	S56229	nitrile reductase
539	16	66.7	1044	2	T43155	nitrile reductase
540	16	66.7	1044	2	T43160	nitrile reductase

541	16	66.7	1046	2	A26838	614	15	62.5	27	2	PC4234	hypothetical prote
542	16	66.7	1069	2	S27922	615	15	62.5	31	2	B61014	defensin R-2 - rat
543	16	66.7	1084	2	T15616	616	15	62.5	34	2	C44336	neurotoxin Tx-3 -
544	16	66.7	1099	2	T16283	617	15	62.5	35	2	S18224	filamentous hemag
545	16	66.7	1101	2	T16840	618	15	62.5	40	2	U70515	Ig heavy chain V-I
546	16	66.7	1129	2	A47511	619	15	62.5	42	2	A17907	thyrotropin beta c
547	16	66.7	1132	2	JM0091	620	15	62.5	45	2	AC3539	hypothetical prote
548	16	66.7	1132	2	JC4127	621	15	62.5	46	2	B81072	hypothetical prote
549	16	66.7	1153	2	P84468	622	15	62.5	46	2	CS3613	plectoxin IX - spi
550	16	66.7	1218	2	B64537	623	15	62.5	52	1	XBPI	bromelain inhibito
551	16	66.7	1235	2	C69165	624	15	62.5	52	2	S66609	bromelain inhibito
552	16	66.7	1236	2	T50904	625	15	62.5	52	2	T10299	conotoxin-like pro
553	16	66.7	1274	2	T10729	626	15	62.5	53	2	T10405	conotoxin-like pro
554	16	66.7	1297	2	T52065	627	15	62.5	53	2	C72850	conotoxin homolog
555	16	66.7	1297	2	T30274	628	15	62.5	53	2	T30499	conotoxin-like pro
556	16	66.7	1307	2	G96711	629	15	62.5	54	1	FEMR	ferradoxin 214Fe-4
557	16	66.7	1307	2	G70535	630	15	62.5	56	1	TRPG	pancreatic secreto
558	16	66.7	1316	2	G70535	631	15	62.5	57	1	NILSA	main precursor -
559	16	66.7	1321	2	JB0352	632	15	62.5	58	2	D82759	hypothetical prote
560	16	66.7	1336	2	T23310	633	15	62.5	59	2	T44147	B3 protein [import
561	16	66.7	1347	2	T41321	634	15	62.5	59	2	AE2319	hypothetical prote
562	16	66.7	1357	2	T16860	635	15	62.5	62	2	F82332	bacterioferritin-a
563	16	66.7	1360	2	T33922	636	15	62.5	63	2	S25772	testis-specific pr
564	16	66.7	1373	2	JB0095	637	15	62.5	63	2	A34484	metallothionein I
565	16	66.7	1376	2	S63986	638	15	62.5	64	2	AD2881	hypothetical prote
566	16	66.7	1396	1	VCBE40	639	15	62.5	66	2	T17934	hypothetical prote
567	16	66.7	1413	2	D44481	640	15	62.5	68	2	B84730	hypothetical prote
568	16	66.7	1416	2	B88550	641	15	62.5	69	2	B97230	hypothetical prote
569	16	66.7	1444	2	T18856	642	15	62.5	70	2	S23316	hypothetical prote
570	16	66.7	1469	2	B36665	643	15	62.5	70	2	C83620	hypothetical prote
571	16	66.7	1475	2	A60026	644	15	62.5	71	2	F69981	hypothetical prote
572	16	66.7	1509	2	T19486	645	15	62.5	71	2	T00008	copy number contro
573	16	66.7	1518	2	T28880	646	15	62.5	71	2	A59412	KGD-bearing platel
574	16	66.7	1523	2	T13953	647	15	62.5	71	2	A59413	platelet-aggregati
575	16	66.7	1531	2	T42218	648	15	62.5	71	2	S13168	baroxostatin ba
576	16	66.7	1557	2	T28811	649	15	62.5	71	2	G43019	platelet aggregati
577	16	66.7	1599	2	T16210	650	15	62.5	72	2	A42325	orf 5' to pheC - p
578	16	66.7	1620	2	T27283	651	15	62.5	72	2	AB2528	hypothetical prote
579	16	66.7	1650	2	S53457	652	15	62.5	72	2	F43019	platelet aggregati
580	16	66.7	1700	2	S08167	653	15	62.5	72	2	D43019	platelet aggregati
581	16	66.7	1772	2	A45532	654	15	62.5	72	2	T43019	platelet aggregati
582	16	66.7	1784	2	T02844	655	15	62.5	72	2	B43019	platelet aggregati
583	16	66.7	1801	1	MMRTS	656	15	62.5	72	2	B43020	platelet aggregati
584	16	66.7	1822	2	S63985	657	15	62.5	73	2	T42302	hypothetical prote
585	16	66.7	1896	2	T01490	658	15	62.5	73	2	S78719	protein YER091c-a
586	16	66.7	2022	2	T48818	659	15	62.5	73	2	H43019	platelet aggregati
587	16	66.7	2049	2	T29227	660	15	62.5	73	2	C43019	platelet aggregati
588	16	66.7	2180	2	T29764	661	15	62.5	73	2	E43019	platelet aggregati
589	16	66.7	2182	2	T14320	662	15	62.5	73	2	A43020	platelet aggregati
590	16	66.7	2233	2	T28669	663	15	62.5	73	2	A43019	platelet aggregati
591	16	66.7	2287	2	T21312	664	15	62.5	73	2	B40003	platelet aggregati
592	16	66.7	2399	2	H71879	665	15	62.5	73	2	A40003	platelet aggregati
593	16	66.7	2470	2	I50726	666	15	62.5	74	2	S33833	hypothetical prote
594	16	66.7	2529	2	B64635	667	15	62.5	74	2	H90820	hypothetical prote
595	16	66.7	2531	2	T16743	668	15	62.5	75	2	S19696	mating pheromone E
596	16	66.7	2548	2	B59435	669	15	62.5	76	1	TRAM3	proteinase inhibit
597	16	66.7	2643	2	T28149	670	15	62.5	76	2	I83048	FSH beta-subunit -
598	16	66.7	2704	2	S09118	671	15	62.5	76	2	G69226	hypothetical prote
599	16	66.7	2718	2	A23475	672	15	62.5	76	2	A39034	neurogranin - bovi
600	16	66.7	2761	2	T21064	673	15	62.5	76	2	D44007	apoptoxin IV - tra
601	16	66.7	3133	2	S52093	674	15	62.5	76	2	C44007	apoptoxin VI - tra
602	16	66.7	3672	2	T23433	675	15	62.5	76	2	AH2120	hypothetical prote
603	16	66.7	3704	2	T37316	676	15	62.5	77	2	JN0536	hypothetical 8.3K
604	16	66.7	4056	2	H66599	677	15	62.5	78	2	D84620	hypothetical prote
605	16	66.7	4544	1	S02392	678	15	62.5	78	2	G98261	hypothetical prote
606	16	66.7	4545	1	S25111	679	15	62.5	79	2	C60767	retrovirus-related
607	16	66.7	4548	1	S00657	680	15	62.5	79	2	AH3369	hypothetical prote
608	16	66.7	4660	2	T42737	681	15	62.5	80	2	I65235	testicular luteini
609	16	66.7	4753	1	A47437	682	15	62.5	81	2	C65019	hypothetical prote
610	16	66.7	5107	2	T29144	683	15	62.5	81	2	AC1864	hypothetical prote
611	16	66.7	10797	2	T30192	684	15	62.5	83	1	W8BP67	gene 18.7 protein
612	15	62.5	15	2	B59045	685	15	62.5	83	2	F35962	blatan alpha - putf
613	15	62.5	27	2	S55030	686	15	62.5	83	2	A34156	blitstatin - putf

687	15	62.5	85	1	W9BPB7	gene 19.2 protein	760	15	62.5	119	2	PH1554	Ig H chain V regio
688	15	62.5	85	2	A1Z765	hypothetical prote	761	15	62.5	119	2	A45574	13.3 kDa polypepti
689	15	62.5	86	2	S20471	class V zygote-age	762	15	62.5	120	1	PSKFP3	phospholipase A2 (
690	15	62.5	87	2	T18115	hypothetical prote	763	15	62.5	120	1	PSKFP4	phospholipase A2 (
691	15	62.5	88	2	P88129	protein F52B1.6 (i	764	15	62.5	120	2	I50993	gonadotropin I bet
692	15	62.5	89	2	P90453	ferredoxin like pr	765	15	62.5	120	2	G71105	hypothetical prote
693	15	62.5	91	2	C38180	hypothetical prote	766	15	62.5	120	2	A12181	hypothetical prote
694	15	62.5	92	2	B30856	hypothetical prote	767	15	62.5	121	1	PSBGAC	phospholipase A2 (
695	15	62.5	94	2	T03285	anther-specific pr	768	15	62.5	121	1	RWBEM3	surface glycoprote
696	15	62.5	95	1	BORT3	prostatic steroid-	769	15	62.5	121	2	PH1661	Ig heavy chain V r
697	15	62.5	96	2	A57483	3-mercaptopyruvate	770	15	62.5	122	1	PSABA	phospholipase A2 (
698	15	62.5	97	2	B97832	alpha-(1,3)-fucosy	771	15	62.5	122	1	PSRSAB	phospholipase A2 (
699	15	62.5	98	2	T33302	hypothetical prote	772	15	62.5	122	1	PSRSAB	phospholipase A2 (
700	15	62.5	98	2	S42596	hypothetical prote	773	15	62.5	122	1	PSVII	phospholipase A2 i
701	15	62.5	100	2	G87427	hypothetical prote	774	15	62.5	122	2	A53872	phospholipase A2 (
702	15	62.5	101	1	MMVZ11	hypothetical prote	775	15	62.5	122	2	A25806	phospholipase A2 (
703	15	62.5	101	1	MMVZ12	11k protein - vacc	776	15	62.5	122	2	A44179	phospholipase A2 (
704	15	62.5	101	1	T30793	11k protein - vacc	777	15	62.5	122	2	S46979	phospholipase A2 -
705	15	62.5	101	2	S73700	probable DNA-bind	778	15	62.5	122	2	S62780	phospholipase A2 (
706	15	62.5	102	2	S50530	hypothetical prote	779	15	62.5	122	2	JX0063	phospholipase A2 (
707	15	62.5	102	2	S69884	hypothetical prote	780	15	62.5	122	2	S09314	phospholipase A2 (
708	15	62.5	102	2	C81889	hypothetical prote	781	15	62.5	122	2	T49039	phospholipase A2 (
709	15	62.5	102	2	B82751	hypothetical prote	782	15	62.5	123	2	B53872	response regulator
710	15	62.5	103	2	C69000	hypothetical prote	783	15	62.5	123	1	PSPGA2	phospholipase A2 (
711	15	62.5	103	2	I58936	hypothetical prote	784	15	62.5	124	1	NRCM	pancreatic ribonuc
712	15	62.5	104	2	D71035	hypothetical prote	785	15	62.5	124	1	NRCMB	pancreatic ribonuc
713	15	62.5	104	2	B81283	hypothetical prote	786	15	62.5	124	1	NRCMB	pancreatic ribonuc
714	15	62.5	104	2	S50911	hypothetical prote	787	15	62.5	124	2	A26535	phospholipase A2 (
715	15	62.5	106	2	S74084	foliitropin beta c	788	15	62.5	124	2	A59420	phospholipase A2 (
716	15	62.5	106	2	B69048	conserved hypotnet	789	15	62.5	124	2	I51190	phospholipase A2 -
717	15	62.5	106	2	S32032	Sp17 protein precu	790	15	62.5	124	2	S25113	insulin-like growt
718	15	62.5	108	2	T45063	hypothetical prote	791	15	62.5	125	2	JX0052	phospholipase A2 (
719	15	62.5	109	2	PH1653	Ig heavy chain V r	792	15	62.5	125	2	S65624	phospholipase A(2)
720	15	62.5	109	2	T46171	hypothetical prote	793	15	62.5	125	2	AD0297	probable membrane
721	15	62.5	109	2	B70956	hypothetical prote	794	15	62.5	125	2	T28743	hypothetical prote
722	15	62.5	110	2	S55208	hypothetical prote	795	15	62.5	126	2	I36914	CS9 protein - bab
723	15	62.5	110	1	A29654	proteinaase inhibit	796	15	62.5	126	2	P84240	hypothetical prote
724	15	62.5	111	2	S69911	Ig V-D-J region (R	797	15	62.5	126	2	G90350	conserved hypotnet
725	15	62.5	111	2	H91187	hypothetical prote	798	15	62.5	127	2	S40034	Killer toxin KP4 P
726	15	62.5	111	2	H85046	hypothetical prote	799	15	62.5	127	2	T10546	Rabl1 protein homo
727	15	62.5	112	2	B29654	proteinaase inhibit	800	15	62.5	128	1	RWHU59	surface glycoprote
728	15	62.5	113	1	S07092	gonadotropin beta	801	15	62.5	128	2	I36894	CS9 protein - gre
729	15	62.5	113	2	T36789	hypothetical prote	802	15	62.5	128	2	T35073	probable phosphori
730	15	62.5	115	2	S66924	probable membrane	803	15	62.5	128	2	AH3422	hypothetical membr
731	15	62.5	115	2	T45252	phosphoribosyl-AMP	804	15	62.5	129	1	FTTHUB	foliitropin beta c
732	15	62.5	117	2	A34860	phospholipase A2 (805	15	62.5	129	1	FTTGB	foliitropin beta c
733	15	62.5	117	2	B32416	phospholipase A2 (806	15	62.5	129	2	A23550	foliitropin beta c
734	15	62.5	117	2	A32416	phospholipase A2 (807	15	62.5	129	2	S73514	probable lipoprote
735	15	62.5	117	2	T26924	hypothetical prote	808	15	62.5	130	2	JC4526	foliitropin beta c
736	15	62.5	118	1	PSKFT1	phospholipase A2 (809	15	62.5	130	2	A32893	foliitropin beta c
737	15	62.5	118	1	PSKFT2	phospholipase A2 (810	15	62.5	131	2	I56894	complement regulat
738	15	62.5	118	1	PSKFT3	phospholipase A2 (811	15	62.5	131	2	A46298	pigment deposition
739	15	62.5	118	1	PSOXB	phospholipase A2 h	812	15	62.5	131	2	A83350	hypothetical prote
740	15	62.5	118	1	FTTHOB	foliitropin beta c	813	15	62.5	132	1	PSHOA	phospholipase A2 (
741	15	62.5	118	2	A32622	phospholipase A2 (814	15	62.5	132	2	I37143	agouti protein pre
742	15	62.5	118	2	B34860	phospholipase A2 (815	15	62.5	132	2	B83382	hypothetical prote
743	15	62.5	118	2	C34860	phospholipase A2 (816	15	62.5	133	1	PSOXG	phospholipase A2 (
744	15	62.5	118	2	S01801	phospholipase A2 (817	15	62.5	133	1	T1RZBR	trypsin inhibitor
745	15	62.5	118	2	PN0141	lutropin beta chai	818	15	62.5	133	2	S14728	textilotoxin chain
746	15	62.5	118	2	PN0139	lutropin beta chai	819	15	62.5	134	2	H84868	hypothetical prote
747	15	62.5	118	2	S61051	hypothetical prote	820	15	62.5	134	1	C64872	hypothetical prote
748	15	62.5	118	2	T49515	hypothetical prote	821	15	62.5	135	2	E83461	hypothetical prote
749	15	62.5	118	2	T32397	hypothetical prote	822	15	62.5	135	2	T15610	hypothetical prote
750	15	62.5	119	1	PSNJ2K	phospholipase A2 (823	15	62.5	136	2	AB0846	formate hydrogenly
751	15	62.5	119	1	PSNJ3K	phospholipase A2 (824	15	62.5	136	2	S08626	formate hydrogenly
752	15	62.5	119	1	PSBYA	phospholipase A2 (825	15	62.5	136	2	F91075	formate hydrogenly
753	15	62.5	119	1	PSNJ2W	phospholipase A2 (826	15	62.5	136	2	H82787	hypothetical prote
754	15	62.5	119	1	PSNJ3W	phospholipase A2 (827	15	62.5	136	2	T23996	hypothetical prote
755	15	62.5	119	1	PSOXA	phospholipase A2 (828	15	62.5	136	2	T36241	hypothetical prote
756	15	62.5	119	2	JN0403	phospholipase A2 (829	15	62.5	137	1	JC4877	phospholipase A2 h
757	15	62.5	119	2	S07528	phospholipase A2 (830	15	62.5	137	2	JC4710	beta-bungarotoxin
758	15	62.5	119	2	JX0115	pancreatic ribonac	831	15	62.5	137	2	S37848	hypothetical prote
759	15	62.5	119	2	A61465	lutropin beta chai	832	15	62.5	137	2	T49243	hypothetical prote

833	15	62.5	137	2	T15609	hypothetical prote	906	15	62.5	146	2	148342	phospholipase A2 (
834	15	62.5	137	2	A96914	uncharacterized pr	907	15	62.5	146	2	T01683	immunization ant
835	15	62.5	137	2	A82763	transcription regu	908	15	62.5	147	1	PSKFA4	phospholipase A2 (
836	15	62.5	137	2	AF2790	conserved hypothet	909	15	62.5	147	2	152489	ribonuclease 4 (RC
837	15	62.5	137	2	D97544	transcription regu	910	15	62.5	147	2	AS3180	ribonuclease PL3 (
838	15	62.5	138	1	151380	phospholipase A2 h	911	15	62.5	147	2	150994	gonadotropin II be
839	15	62.5	138	1	PSRSAT	phospholipase A2 h	912	15	62.5	147	2	T20563	hypothetical prote
840	15	62.5	138	1	PSRVIF	phospholipase A2	913	15	62.5	147	2	E82523	hypothetical prote
841	15	62.5	138	1	PSVIAA	phospholipase A2 (914	15	62.5	147	2	F70216	hypothetical prote
842	15	62.5	138	1	PSVIAC	phospholipase A2 (915	15	62.5	148	2	H82065	PTS system, nitrog
843	15	62.5	138	1	BGMU2	spermatid transiti	916	15	62.5	148	2	147180	Ig heavy chain var
844	15	62.5	138	2	A46188	phospholipase A2 (917	15	62.5	149	2	T01676	cytochrome-c oxida
845	15	62.5	138	2	A46169	phospholipase A2 (918	15	62.5	149	2	S70979	bDPL protein - Esc
846	15	62.5	138	2	F46188	phospholipase A2 (919	15	62.5	149	2	AG2708	conserved hypothet
847	15	62.5	138	2	S29299	phospholipase A2 (920	15	62.5	150	2	F69883	conserved hypothet
848	15	62.5	138	2	S59522	phospholipase A2 (921	15	62.5	151	2	C90048	conserved hypothet
849	15	62.5	138	2	S10333	ammodoroxin B prec	922	15	62.5	152	2	G81161	hypothetical prote
850	15	62.5	138	2	151386	phospholipase A2 (923	15	62.5	153	2	T19054	hypothetical prote
851	15	62.5	138	2	S10992	ammodoroxin L precu	924	15	62.5	153	2	PN0103	hypothetical 17K p
852	15	62.5	138	2	JC1342	phospholipase A2 (925	15	62.5	154	2	S35207	proteinnase 7 - buf
853	15	62.5	138	2	150098	phospholipase a2 -	926	15	62.5	154	2	S23320	hypothetical prote
854	15	62.5	138	2	A29479	thymotropin beta c	927	15	62.5	154	2	T27967	hypothetical prote
855	15	62.5	138	2	PC1211	envelope protein -	928	15	62.5	156	2	T11839	hypothetical prote
856	15	62.5	138	2	PC1212	envelope protein -	929	15	62.5	156	2	S74049	hypothetical prote
857	15	62.5	138	2	PC1205	envelope protein -	930	15	62.5	157	2	G65156	putative electron
858	15	62.5	138	2	PC1203	envelope protein -	931	15	62.5	157	2	S58017	probable olfactory
859	15	62.5	138	2	PC1200	genome polypotein	932	15	62.5	157	2	S58338	hypothetical prote
860	15	62.5	138	2	PC1210	envelope protei	933	15	62.5	158	2	T52365	hypothetical prote
861	15	62.5	138	2	PC1207	envelope protei	934	15	62.5	158	2	G86964	hypothetical prote
862	15	62.5	138	2	PC1197	genome polypotein	935	15	62.5	158	2	T22829	hypothetical prote
863	15	62.5	138	2	PC1209	envelope protei	936	15	62.5	159	2	S46822	hypothetical prote
864	15	62.5	138	2	PC1199	genome polypotein	937	15	62.5	159	2	H81657	conserved hypothet
865	15	62.5	138	2	PC1201	envelope protei	938	15	62.5	159	2	D71500	hypothetical prote
866	15	62.5	138	2	T14185	chitinase (EC 3.2.	939	15	62.5	159	2	T33696	hypothetical prote
867	15	62.5	138	2	T21792	hypothetical prote	940	15	62.5	160	2	T23995	hypothetical prote
868	15	62.5	139	1	PSRV	phospholipase A2 (941	15	62.5	161	2	H64504	hypothetical prote
869	15	62.5	139	2	152320	testicular luteini	942	15	62.5	161	2	S40711	hypothetical prote
870	15	62.5	139	2	A38612	insulin-like growt	943	15	62.5	161	2	T48146	hypothetical prote
871	15	62.5	139	2	S31612	beta-1,3-glucanase	944	15	62.5	162	2	H81058	disulfide bond for
872	15	62.5	139	2	AG0578	conserved hypothet	945	15	62.5	162	2	T50253	Vacuolar ATP synth
873	15	62.5	140	1	11MSG1	interleukin-4 prec	946	15	62.5	162	2	T18188	hypothetical prote
874	15	62.5	140	1	HBFGC	hemoglobin beta ch	947	15	62.5	162	2	A72738	hypothetical prote
875	15	62.5	140	2	A46166	gonadotropin II be	948	15	62.5	162	2	B86387	18.7K hypothetica
876	15	62.5	140	2	H69405	hypothetical prote	949	15	62.5	163	2	A71469	hypothetical prote
877	15	62.5	140	2	T29914	hypothetical prote	950	15	62.5	164	1	A91016	ferredoxin-type pr
878	15	62.5	140	2	E87035	hypothetical prote	951	15	62.5	164	1	C95860	ferredoxin-type pr
879	15	62.5	141	1	B60626	gonadotropin beta	952	15	62.5	164	1	F64990	ferredoxin-type pr
880	15	62.5	141	1	UTB0B	lutropin beta chai	953	15	62.5	164	2	C35216	tumor suppressor p
881	15	62.5	141	1	UTB0B	lutropin beta chai	954	15	62.5	164	2	184725	FP18 protein - tow
882	15	62.5	141	1	UTRFB	lutropin beta chai	955	15	62.5	165	1	KTRAB	choriogonadotropin
883	15	62.5	141	1	UTRFB	lutropin beta chai	956	15	62.5	165	1	KTRAB	peptidoglycan asso
884	15	62.5	141	1	UTSHB	lutropin beta chai	957	15	62.5	165	2	140769	hypothetical prote
885	15	62.5	141	2	UT4527	luteinizing hormon	958	15	62.5	165	2	H84634	hypothetical prote
886	15	62.5	142	2	S54243	Ig mu heavy chain	959	15	62.5	165	2	E84346	hypothetical prote
887	15	62.5	142	2	AD2696	11pA protein [impo	960	15	62.5	165	2	S59899	chlorocornurin chai
888	15	62.5	142	2	B36711	F24J5.7 [imported]	961	15	62.5	166	2	JN0248	platelet-derived g
889	15	62.5	143	2	T20951	gonadotropin prote	962	15	62.5	166	2	JC1348	hypothetical 18K p
890	15	62.5	144	1	UTCAB	gonadotropin beta	963	15	62.5	166	2	A40783	18.2K protein - ph
891	15	62.5	144	1	UT0089	interleukin-9 prec	964	15	62.5	170	2	C96009	probable plasmid s
892	15	62.5	144	2	E88485	protein F23F12.3 (965	15	62.5	171	2	A10607	probable lipoprote
893	15	62.5	144	2	T16824	hypothetical prote	966	15	62.5	171	2	D90747	probable enzyme (i
894	15	62.5	145	1	PSBOA	phospholipase A2 (967	15	62.5	171	2	A64825	ybip protein precu
895	15	62.5	145	1	PSKFA2	phospholipase A2 (968	15	62.5	171	2	T13072	hypothetical prote
896	15	62.5	145	2	148093	phospholipase A2 (969	15	62.5	171	2	E85046	hypothetical prote
897	15	62.5	145	2	137231	beta-gonadotropin	970	15	62.5	172	2	T20210	hypothetical prote
898	15	62.5	146	1	PSD6	phospholipase A2 (971	15	62.5	172	2	T10346	hypothetical prote
899	15	62.5	146	1	PSNJAF	phospholipase A2 (972	15	62.5	173	1	B45932	actrial gland pepti
900	15	62.5	146	1	PSPGA	phospholipase A2 (973	15	62.5	173	1	COGAAA	actrial gland pepti
901	15	62.5	146	1	S34049	phospholipase A2 (974	15	62.5	173	1	GOGABA	actrial gland pepti
902	15	62.5	146	1	S16763	gonadotropin beta	975	15	62.5	173	2	S62349	L71-3 protein - fr
903	15	62.5	146	2	A35493	phospholipase A2 (976	15	62.5	173	2	H85597	probable enzyme yb
904	15	62.5	146	2	A33394	phospholipase A2 (977	15	62.5	173	2	E97569	hypothetical prote
905	15	62.5	146	2	JU0283	phospholipase A2 (978	15	62.5	174	2	T15176	hypothetical prote

979 15 62.5 174 2 A86358 Similar to blue co
980 15 62.5 175 2 T48541 hypothetical prote
981 15 62.5 176 2 T18921 hypothetical prote
982 15 62.5 177 1 AFMD8 phycoerythrin beta
983 15 62.5 177 1 A55537 CDdiacylglycerol-
984 15 62.5 177 2 C71329 hypothetical prote
985 15 62.5 178 2 T29345 hypothetical prote
986 15 62.5 178 2 C84723 probable PHD-type
987 15 62.5 178 2 D75390 NADH dehydrogenase
988 15 62.5 180 2 S14019 hypothetical prote
989 15 62.5 180 2 B84768 hypothetical prote
990 15 62.5 180 2 F75607 areneate reductase
991 15 62.5 181 2 D86518 hypothetical prote
992 15 62.5 181 2 C72104 hypothetical prote
993 15 62.5 182 2 T45302 hypothetical prote
994 15 62.5 182 2 T11906 NADH2 dehydrogenas
995 15 62.5 183 2 A85561 hypothetical prote
996 15 62.5 183 2 D64795 Apo-citrate lyase
997 15 62.5 183 2 E90710 Apo-citrate lyase
998 15 62.5 184 2 AG0736 hypothetical prote
999 15 62.5 185 1 JC2394 phospholipase A2 1
1000 15 62.5 185 2 C96808 protein F28K19.4 [

ALIGNMENTS

RESULT 1

Unknown protein, 30275-30808 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C/Accession: F96719

R/Theologian, A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000

C/A: Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maltli, R.; Marzilli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: F96719

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-177 <STO>

A/Cross-references: UNIPROT:Q9C9L7; UNIPARC:UPI00000A1912; GB:AE005173; NID:g6665540; PI

C/Genetics:

A/Map position: 1

Query Match 75.0%; Score 18; DB 2; Length 177;
Best Local Similarity 25.0%; Pred. No. 66;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
Db 87 CSSSSSSC 94

RESULT 2

rhomoid protein - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster

C/Date: 06-Jul-1990 #sequence_revision 06-Jul-1990 #text_change 05-Oct-2004

C/Accession: A34597

R/Bier, E.; Jan, L.Y.; Jan, Y.N.

Genes Dev. 4, 190-203, 1990

A/Title: rhomoid, a gene required for dorsoventral axis establishment and peripheral ne

A/Reference number: A34597; MUID:90249726; PMID:2110920

A/Accession: A34597

A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-355 <BIE>
A/Cross-references: UNIPROT:P20350; UNIPARC:UPI000016BD27; GB:X52454; NID:g10878; PIDN:C
C/Genetics:
A/Map position: 1

Query Match 75.0%; Score 18; DB 2; Length 355;
Best Local Similarity 25.0%; Pred. No. 82;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
Db 57 CSTASSC 64

RESULT 3

TPA-induced protein 11D - mouse

C/Species: Mus musculus (house mouse)

C/Date: 30-Aug-1991 #sequence_revision 30-Aug-1991 #text_change 09-Jul-2004

C/Accession: C39590

R/Varnum, B.C.; Ma, O.; Chi, T.; Fletcher, B.; Herschman, H.R.
Mol. Cell. Biol. 11, 1754-1758, 1991

A/Title: The TIS11 primary response gene is a member of a gene family that encodes prote

A/Reference number: A39590; MUID:91141531; PMID:1956120

A/Accession: C39590

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-367 <VAR>
A/Cross-references: UNIPROT:P23949; UNIPARC:UPI000002A130; GB:M58564; NID:g202066; PIDN:J
C/Keywords: DNA binding; zinc finger

Query Match 75.0%; Score 18; DB 2; Length 367;
Best Local Similarity 25.0%; Pred. No. 82;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
Db 273 CSSSASSC 280

RESULT 4

ERF-2 protein - human

C/Species: Homo sapiens (man)

C/Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004

C/Accession: S49147; 137372

R/Nie, X.F.; Maclean, R.N.; Kumar, V.; McKay, I.A.; Bustin, S.A.

submitted to the EMBL Data Library, April 1994

A/Reference number: S49147

A/Accession: S49147

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-492 <NIE>

A/Cross-references: UNIPROT:P47974; UNIPARC:UPI000003794E

R/Nie, X.F.; Maclean, R.N.; Kumar, V.; McKay, I.A.; Bustin, S.A.

Gene 152, 285-286, 1995

A/Title: ERF-2, the human homologue of the murine Tis11d early response gene.

A/Reference number: 137372; MUID:95137407; PMID:7835719

A/Accession: 137372

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-492 <RES>

A/Cross-references: UNIPARC:UPI000003794E; EMBL:X78992; NID:g509777; PIDN:CAA55592.1; PII

C/Genetics:

A/Map position: 13q12.3-13q12.3

A/Reference number: 13q12.3-13q12.3

A/Accession: 13q12.3-13q12.3

Query Match 75.0%; Score 18; DB 2; Length 492;
Best Local Similarity 25.0%; Pred. No. 90;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
DB 299 CASSSSSC 306

RESULT 5

JDUVL1
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adw, strain Japan/

C/Species: hepatitis B virus, HBV

C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004

C/Accession: D28925

R/Okamoto, H.; Tsuda, F.; Sakugawa, H.; Sastroseowinjo, R.I.; Imai, M.; Miyakawa, Y.; M

J. Gen. Virol. 69, 2575-2583, 1988

A/Title: Typing hepatitis B virus by homology in nucleotide sequence: comparison of surF

A/Reference number: J50253; PMID:89010694; PMID:3171552

A/Accession: D28925

A/Molecule type: DNA

A/Residues: 1-843 <OKA>

A/Cross-references: UNIPROT:P17394; UNIPARC:UPI00001297F9; GB:D00329; NID:G221497

C/Superfamily: hepatitis virus DNA-directed DNA polymerase

C/Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 75.0%; Score 18; DB 1; Length 843;
Best Local Similarity 25.0%; Pred. No. 1.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
DB 268 CASSSSSC 275

RESULT 6

JDUVL2
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adw, strain Okinaw

C/Species: hepatitis B virus, HBV

C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004

C/Accession: E28925

R/Okamoto, H.; Tsuda, F.; Sakugawa, H.; Sastroseowinjo, R.I.; Imai, M.; Miyakawa, Y.; M

J. Gen. Virol. 69, 2575-2583, 1988

A/Title: Typing hepatitis B virus by homology in nucleotide sequence: comparison of surF

A/Reference number: J50253; PMID:89010694; PMID:3171552

A/Accession: E28925

A/Molecule type: DNA

A/Residues: 1-843 <OKA>

A/Cross-references: UNIPROT:P17395; UNIPARC:UPI00001297FC; GB:D00330; NID:G221498

C/Superfamily: hepatitis virus DNA-directed DNA polymerase

C/Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 75.0%; Score 18; DB 1; Length 843;
Best Local Similarity 25.0%; Pred. No. 1.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
DB 268 CASSSSSC 275

RESULT 7

T26859
hypothenical protein Y43F8B.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T26859

R/Almsough, R.

submitted to the EMBL Data Library, October 1998

A/Accession: T26859

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1-1743 <MIL>
A/Cross-references: UNIPROT:Q9XWXS; UNIPARC:UPI0000078FD6; EMBL:AL032623; PIDN:CAA21511.

A/Experimental source: clone Y43F8B

C/Genetics: ...

A/Gene: CESP:Y43F8B.3

A/Introns: 65/1; 92/2; 128/1; 229/1; 367/1; 422/1; 486/2; 523/1; 571/1; 628/1; 857/2; 94

Query Match 75.0%; Score 18; DB 2; Length 1743;
Best Local Similarity 25.0%; Pred. No. 1.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
DB 1123 CSATATSC 1130

RESULT 8

T31687
surface antigen - Paramoecium primaurelia

C/Species: Paramoecium primaurelia

C/Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004

C/Accession: T31687

R/Bourgain-Guglielmetti, F.; Caron, F.

Journal of Eukaryot. Microbiol. 43, 303-314, 1996

A/Title: Molecular characterization of the D surface protein gene subfamily in Paramoeciu

A/Reference number: Z21061; PMID:96313351; PMID:8768434

A/Accession: T31687

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-2543 <BOU>

A/Cross-references: UNIPROT:P90649; UNIPARC:UPI0000079A03; EMBL:X96616; NID:G1235576; PI

C/Genetics: ...

A/Genetic code: SGC5

C/Superfamily: G surface protein

Query Match 75.0%; Score 18; DB 2; Length 2543;
Best Local Similarity 25.0%; Pred. No. 1.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
DB 1961 CTSSTTC 1968

RESULT 9

A44036
collagen alpha 1(XII) chain - bovine (fragment)

C/Species: Bos primigenius taurus (cattle)

C/Date: 27-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C/Accession: A44036

R/Instrum, G.P.; McDonough, A.M.; Martinovitch, M.P.; Keene, D.R.; Morris, N.P.; Burgeso

J. Biol. Chem. 267, 20087-20092, 1992

A/Title: Identification and partial purification of a large, variant form of type XII co

A/Reference number: A44036; PMID:93015874; PMID:1400326

A/Accession: A44036

A/Status: preliminary

A/Molecule type: protein

A/Residues: 1-26 <LUN>

A/Cross-references: UNIPROT:Q9TRM6; UNIPARC:UPI0000087EAB

A/Experimental source: skin

A/Note: sequence extracted from NCBI backbone (NCBIP:115795)

Query Match 70.8%; Score 17; DB 2; Length 26;

Best Local Similarity 25.0%; Pred. No. 89;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
DB 10 CEFATSSC 17

RESULT 10

S72742

B1177 P1 32 protein - Mycobacterium leprae

C:Species: Mycobacterium leprae
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S72742
R:Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A:Description: Mycobacterium leprae cosmid B1177.
A:Reference number: S72694
A:Accession: S72742
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-89 <SM1>
A:Cross-references: UNIPROT:Q49652; UNIPARC:UPI00000B6610; EMBL:U00011; NID:g466807; PID
C:Genetics:
A:Start codon: GTG

Query Match 70.8%; Score 17; DB 2; Length 89;
Best Local Similarity 25.0%; Pred. No. 1.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
Db 71 CRTSSTAC 78

RESULT 11

T42262 hypothetical protein - phage SPPI

C:Species: phage SPPI
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T42262
R:Alonso, J.C.; Iudler, G.; Stiege, A.C.; Chal, S.; Weise, F.; Trautner, T.A.
Gene 204, 201-212, 1997
A:Title: The complete nucleotide sequence and functional organization of Bacillus subtilis
A:Reference number: Z22137; MUID:98094274; PMID:9434185
A:Accession: T42262
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-101 <ALO>
A:Cross-references: UNIPROT:Q48435; UNIPARC:UPI000009B544; EMBL:X97918; PIDN:CAA66572.1

Query Match 70.8%; Score 17; DB 2; Length 101;
Best Local Similarity 25.0%; Pred. No. 1.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
Db 59 CSKTSSAC 66

RESULT 12

S72979 probable L-a-aminoadipyl-L-cystenyl-D-valine synthetase acvs - Mycobacterium leprae

N:Alternate names: B229 C2 179 protein
C:Species: Mycobacterium leprae
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S72979
R:Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A:Description: Mycobacterium leprae cosmid B229.
A:Reference number: S72588
A:Accession: S72979
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-120 <SM1>
A:Cross-references: UNIPROT:Q49859; UNIPARC:UPI00000BE2D7; EMBL:U00020; NID:g467102; PID
C:Genetics:
A:Gene: acvs
A:Start codon: GTG

Query Match 70.8%; Score 17; DB 2; Length 120;
Best Local Similarity 25.0%; Pred. No. 1.4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
Db 4 CSYTSASC 11

RESULT 13

J00150 hypothetical 13k protein - Pseudomonas aeruginosa

C:Species: Pseudomonas aeruginosa
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C:Accession: J00150
R:Kato, J.; Chu, L.; Kitano, K.; DeVault, J.D.; Kimbara, K.; Chakrabarty, A.M.; Misra, T.
Gene 84, 31-38, 1989
A:Title: Nucleotide sequence of a regulatory region controlling alginate synthesis in P.
A:Reference number: J00132; MUID:90108714; PMID:2514124
A:Accession: J00150
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-122 <KAT>
A:Cross-references: UNIPARC:UPI00001791D0
C:Superfamily: Pseudomonas aeruginosa hypothetical 13k protein

Query Match 70.8%; Score 17; DB 2; Length 122;
Best Local Similarity 25.0%; Pred. No. 1.4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
Db 49 CSTATVAC 56

RESULT 14

S25574 Ig heavy chain V region - human

C:Species: Homo sapiens (man)
C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 20-Jun-2000
C:Accession: S25574
R:Adderson, E.B.; Shackelford, P.G.; Quinn, A.; Carroll, W.L.
J. Immunol. 147, 1667-1674, 1991
A:Title: Restrictcd Ig H chain V gene usage in the human antibody response to Haemophilus
A:Reference number: S25571; MUID:9134592; PMID:1908880
A:Accession: S25574
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-123 <ADD>
A:Cross-references: UNIPARC:UPI0000115EBF; EMBL:X56527; NID:G33569; PIDN:CAA39874.1; PID
A>Note: the authors translated the codon GAA for residue 69 as Asp and TTA for residue 10
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:23-108/Domain: immunoglobulin homology <IMW>

Query Match 70.8%; Score 17; DB 2; Length 123;
Best Local Similarity 25.0%; Pred. No. 1.4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
Db 113 CASSAYAC 120

RESULT 15

RKHUY T-cell receptor beta chain precursor V region (YT35) - human

N:Alternate names: T-cell receptor beta-1 chain J-B1.2 segment
C:Species: Homo sapiens (man)
C:Date: 03-Aug-1984 #sequence_revision 03-Aug-1984 #text_change 09-Jul-2004
C:Accession: A02000; E24687
R:Yanagi, Y.; Yoshikai, Y.; Leggett, K.; Clark, S.P.; Aleksander, I.; Mak, T.W.
Nature 308, 145-149, 1984
A:Title: A human T cell-specific cDNA clone encodes a protein having extensive homology to
A:Reference number: A93324; MUID:84142269; PMID:6336315
A:Accession: A02000

A:Molecule type: mRNA
 A:Residues: 1-135 <YAN>
 A:Cross-references: UNIPROT:P01733; UNIPARC:UPI00001377D3; GB:K01571
 A:Experimental source: clone YT35
 R:Toyonaga, B.; Yoshikawa, Y.; Vadasz, V.; Chin, B.; Mak, T. W.
 Proc. Natl. Acad. Sci. U.S.A. 82, 8624-8628, 1985
 A:Title: Organization and sequences of the diversity, joining, and constant region genes
 A:Reference number: A94081; MUID:86094276; PMID:3866244
 A:Accession: E24687
 A:Molecule type: DNA
 A:Residues: 121-135 <TOY>
 A:Cross-references: UNIPARC:UPI0000113C76; GB:M4158; NID:g338844; PIDN:AAA60669.1; PID:
 C:Genetics:
 A:Gene: GDB:TCRB
 A:Cross-references: GDB:120405; OMIM:186930
 A:Map position: 7q35-7q35
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: glycoprotein; heterotetramer; receptor; T-cell
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-135/Product: T-cell receptor beta chain V region YT35 #status predicted <MAT>
 F:35-113/Domain: immunoglobulin homology <IMM>
 F:42-111/Disulfide bonds: #status predicted
 F:90/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 70.8%; Score 17; DB 1; Length 135;
 Best Local Similarity 25.0%; Pred. No. 1.5e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
 Db 111 CASSFTC 118

RESULT 16
 A71123
 hypothetical protein PH0752 - *Pyrococcus horikoshii*
 C:Species: *Pyrococcus horikoshii*
 C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
 C:Accession: A71123
 R:Kawabata, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
 M.; Ohnuki, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
 DNA Res. 5, 55-76, 1998
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
 A:Reference number: A71000; MUID:98344137; PMID:9679194
 A:Accession: A71123
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-139 <KAW>
 A:Cross-references: UNIPROT:O58508; UNIPARC:UPI0000628CE; GB:AP000003; NID:g3236130; PI
 A:Experimental source: strain OT3
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C:Genetics:
 A:Gene: PH0752

Query Match 70.8%; Score 17; DB 2; Length 139;
 Best Local Similarity 25.0%; Pred. No. 1.5e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
 Db 81 CSSTYASC 88

RESULT 17
 H87436
 hypothetical protein CC1513 [imported] - *Caulobacter crescentus*
 C:Species: *Caulobacter crescentus*
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C:Accession: H87436
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Land, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 u, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: H87436
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-145 <STD>
 A:Cross-references: UNIPROT:Q9A851; UNIPARC:UPI00000C73F0; GB:AE005673; NID:g13422892; P
 C:Genetics:
 A:Gene: CC1513

Query Match 70.8%; Score 17; DB 2; Length 145;
 Best Local Similarity 25.0%; Pred. No. 1.5e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
 Db 10 CAABEAC 17

RESULT 18
 T22305
 hypothetical protein F46C5.2 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T22305
 R:Thomas, K.
 submitted to the EMBL Data Library, September 1995
 A:Reference number: Z19544
 A:Accession: T22305
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-182 <WIL>
 A:Cross-references: UNIPROT:P52881; UNIPARC:UPI000013A06E; EMBL:Z54281; PIDN:CAA91050.1;
 A:Experimental source: clone F46C5
 C:Genetics:
 A:Gene: CESP:F46C5.2
 A:Map position: 2
 A:Introns: 14/3; 50/1; 76/3; 145/2

Query Match 70.8%; Score 17; DB 2; Length 182;
 Best Local Similarity 25.0%; Pred. No. 1.6e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
 Db 35 CTSGATTTC 42

RESULT 19
 E70641
 hypothetical protein Rv0698 - *Mycobacterium tuberculosis* (strain H37RV)
 C:Species: *Mycobacterium tuberculosis*
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C:Accession: E70641
 R:Cole, S.T.; Broese, R.; Parkhill, J.; Garlier, T.; Churcher, C.; Harris, D.; Gordon, S.
 Conroy, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandram, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: E70641
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-203 <COL>
 A:Cross-references: UNIPROT:P95044; UNIPARC:UPI00000C14B5; GB:Z84395; GB:AL123456; NID:g
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: Rv0698

Query Match 70.8%; Score 17; DB 2; Length 203;
 Best Local Similarity 25.0%; Pred. No. 1.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
Db 51 CTTAATTC 58

RESULT 20

H75277
hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: H75277

R:White, O.; Eksen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: H75277

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-212 <WHI>

A:Cross-references: UNIPROT:Q9RKS8; UNIPARC:UPI00000CIAF7; GB:AE002071; GB:AE000513; NID

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR2407

A:Map position: 1

Query Match 70.8%; Score 17; DB 2; Length 212;
Best Local Similarity 25.0%; Pred. No. 1.7e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
Db 15 CRSASAC 22

RESULT 21

A41128

myogenin - human

N:Alternate names: myogenic determination factor Myf-4

C:Species: Homo sapiens (man)

C>Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 09-Jul-2004

C:Accession: A41128; S27295; S06948; S21978

R:Salminen, A.; Braun, T.; Buchberger, A.; Ueters, S.; Winter, B.; Arnold, H.H.

J. Cell Biol. 115, 905-917, 1991

A:Title: Transcription of the muscle regulatory gene MYF4 is regulated by serum component

A:Reference number: A41128; MUID:92064650; PMID:1659574

A:Accession: A41128

A:Molecule type: DNA

A:Residues: 1-224 <SML>

A:Cross-references: UNIPROT:P15173; UNIPARC:UPI000012FBA4; EMBL:X62155

R:Braun, T.; Bober, E.; Buschhausen-Denker, G.; Kohz, S.; Grzeschik, K.H.; Arnold, H.H.

EMBO J. 9, 592, 1990

A:Reference number: S27295

A:Contents: extratum

A:Accession: S27295

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-224

A:Cross-references: UNIPARC:UPI000012FBA4

A>Note: this is a revision to the sequence from reference S06947

R:Braun, T.; Bober, E.; Buschhausen-Denker, G.; Kotz, S.; Grzeschik, K.H.; Arnold, H.H.

EMBO J. 8, 3617-3625, 1989

A:Title: Differential expression of myogenic determination genes in muscle cells: possib

A:Reference number: S06947; MUID:90059960; PMID:2583111

A:Accession: S06948

A:Molecule type: mRNA

A:Residues: 1-154; 'SOGCPANALATAPAVOSGAVHWSAPTOGIIISRLTLQWPTCTSPSPWTASQWKCIMPQMKPC

A:Cross-references: UNIPARC:UPI000016AD95; EMBL:X17651; NID:934831; PIDN:CAA35641.1; PID

A>Note: this sequence has been revised in reference S27295

R:Arnold, H.H.

submitted to the EMBL Data Library, September 1991

A:Reference number: S21978

A:Accession: S21978

A:Molecule type: DNA

A:Residues: 1-223 <ARN>

A:Cross-references: UNIPARC:UPI000016AD96; EMBL:X62155; NID:934833; PIDN:CAA44080.1; PID

C:Genetics:

A:Gene: GDB:MYOG; MYF4

A:Cross-references: GDB:120210; OMIM:159980

A:Map position: 1q31-1q41

C:Introns: 157/3; 184/3

C:Superfamily: human myogenin

C:Keywords: DNA binding

Query Match 70.8%; Score 17; DB 2; Length 224;
Best Local Similarity 25.0%; Pred. No. 1.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
Db 162 CSHSASC 169

RESULT 22

S09778

hypothetical protein UL16 precursor - human cytomegalovirus (strain AD169)

C:Species: human cytomegalovirus, human herpesvirus 5

A>Note: host Homo sapiens (man)

C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004

C:Accession: S09778

R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; F

M.; Barrell, B.G.

curr. Top. Microbiol. Immunol. 154, 125-169, 1990

A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus

A:Reference number: S09749; MUID:90269039; PMID:2161319

A:Accession: S09778

A:Molecule type: DNA

A:Residues: 1-230 <CHE>

A:Cross-references: UNIPROT:P16757; UNIPARC:UPI0000137B56; EMBL:X17403; NID:959591; PIDN

A>Note: this sequence was submitted to the EMBL Data Library, December 1989

C:Superfamily: human cytomegalovirus hypothetical protein UL16

C:Keywords: glycoprotein; transmembrane protein

F:1-55/Domain: signal sequence #status predicted <SIG>

F:126-230/Product: hypothetical protein UL16 #status predicted <MAT>

F:189-206/Domain: transmembrane #status predicted <TM>

F:35; 41; 68; 84; 95; 101; 132; 145/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 70.8%; Score 17; DB 2; Length 230;
Best Local Similarity 25.0%; Pred. No. 1.7e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
Db 18 CLSASSC 25

RESULT 23

F95411

hypothetical protein Sma2221 [imported] - Sinorhizobium meliloti (strain 1021) megaplasm

C:Species: Sinorhizobium meliloti

C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004

C:Accession: F95411

R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: F95411

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-237 <KIR>

A:Cross-references: UNIPROT:Q92XP6; UNIPARC:UPI00000CB344; GB:AE006469; PIDN:AAK55856.1;

A:Experimental source: strain 1021, megaplasmid pSymA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, M.; Vandenbol, M.; Vorholter, F.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.
 A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 C:Contents: annotation
 C:Genetics:
 A:Gene: Sma2221
 A:Genome: plasmid

Query Match 70.8%; Score 17; DB 2; Length 237;
 Best Local Similarity 25.0%; Pred. No. 1.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
 Db 76 CTAATTSC 83

RESULT 24

JH0367
 ribonuclease (EC 3.1.27.-) Rh precursor - *Rhizopus niveus*
 N:Alternate names: base nonspecific ribonuclease
 C:Species: *Rhizopus niveus*

C:Date: 05-Mar-1999 #sequence_revision 31-Jan-1997 #text_change 31-Dec-2004

C:Accession: JH0367; A1449; B4149

R:Ohji, K.; Horinuchi, H.; Watanabe, H.; Takagi, M.; Yano, K.; Irie, M.

J. Biochem. 109, 776-785, 1991
 A:Title: Expression of RNase Rh from *Rhizopus niveus* in yeast and characterization of the

A:Reference number: JH0367; MUID:92011465; PMID:1655721

A:Accession: JH0367

A:Molecule type: DNA

A:Residues: 1-238 <OHG>
 A:Cross-references: UNIPROT:P08056; UNIPARC:UPI00001344AC; DBJ:D12476; DBJ:D01125; NID

J. Horinuchi, H.; Yano, K.; Takagi, M.; Yano, K.; Wakabayashi, E.; Sada, A.; Mane, S.; C

J. Biochem. 103, 408-418, 1988

A:Title: Primary structure of a base non-specific ribonuclease from *Rhizopus niveus*.

A:Reference number: A1449; MUID:88273061; PMID:3391995

A:Accession: A1449

A:Molecule type: DNA

A:Residues: 1-81, 'SLV', 85-238 <HOR>
 A:Cross-references: UNIPARC:UPI000016897A; GB:D00238; NID:G218044; PIDN:BA00167.1; PID:

A>Note: the authors translated the codon TCA for residue 82 as Asn, CTC for residue 83 &

A:Accession: B4149

A:Molecule type: protein

A:Residues: 17-238 <HO2>

A:Cross-references: UNIPARC:UPI000011099A
 R:Kuribara, H.; Nomura, T.; Mitani, Y.; Ohgi, K.; Irie, M.; Nakamura, K.T.

J. Mol. Biol. 255, 310-320, 1996

A:Title: The crystal structure of ribonuclease Rh from *Rhizopus niveus* at 2.0 Å resolution

A:Reference number: S62310; MUID:96150305; PMID:8551522

C:Contents: annotation; X-ray crystallography, 2.0 angstroms

C:Genetics:

A:Introns: 21/1; 116/2; 151/3; 167/2

C:Function:
 A:Description: hydrolyzes internal phosphodiester bonds of RNA to produce 3'-phosphomono

C:Superfamily: RNases

C:Keywords: hydrolase

F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-238/Product: ribonuclease Rh #status experimental <MAT>
 F:19-36; 26-69; 75-136; 79-128; 198-229/Disulfide bonds: #status experimental
 F:62; 121; 125/Active site: His, Glu, His #status predicted
 F:65; 120; 124/Binding site: Substrate (Trp, His, Lys) #status predicted

Query Match 70.8%; Score 17; DB 1; Length 238;
 Best Local Similarity 25.0%; Pred. No. 1.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8

Db 19 CSSTALSC 26

RESULT 25

A35871
 trypsin (EC 3.4.21.4) precursor, pancreatic - African clawed frog
 C:Species: *Xenopus laevis* (African clawed frog)

C:Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 09-Jul-2004

C:Accession: A35871; S12117

R:Shi, Y.B.; Brown, D.D.

Genes Dev. 4, 1107-1113, 1990

A:Title: Developmental and thyroid hormone-dependent regulation of pancreatic genes in X

A:Reference number: A35871; MUID:91007255; PMID:2210372

A:Accession: A35871

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-243 <SHI>
 A:Cross-references: UNIPROT:P19799; UNIPARC:UPI00001376E0; EMBL:X53458; NID:965162; PIDN

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; protein digestion; serine proteinase

F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-20/Domain: activation peptide #status predicted <APT>
 F:21-236/Domain: trypsin homology <TRY>
 F:26-243/Product: trypsin I #status predicted <MAT>
 F:27-157; 45-61; 129-230; 168-182/Disulfide bonds: #status predicted
 F:60; 104; 197/Active site: His, Asp, Ser #status predicted

Query Match 70.8%; Score 17; DB 2; Length 243;
 Best Local Similarity 25.0%; Pred. No. 1.8e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
 Db 129 CSAAGTSC 136

RESULT 26

T47902
 hypothetical protein T4C21.320 - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004

C:Accession: T47902

R:Choisme, N.; Robert, C.; Brothier, P.; Catroliano, L.; Artiguenave, F.; San

.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24479

A:Accession: T47902

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-248 <CHO>

A:Cross-references: UNIPROT:Q9LZX3; UNIPARC:UPI000009C7A5; EMBL:AL162295

A:Experimental source: cultivar Columbia; BAC clone T4C21

C:Genetics:

A:Map position: 3

A:Introns: 61/3; 96/1; 119/1; 125/3; 143/2; 153/3; 181/1

A>Note: T4C21.320

Query Match 70.8%; Score 17; DB 2; Length 248;
 Best Local Similarity 25.0%; Pred. No. 1.8e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
 Db 191 CSSAKSC 198

RESULT 27

E87269
 exopolysaccharide production protein Pss [imported] - *Caulobacter crescentus*
 C:Species: *Caulobacter crescentus*

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 31-Dec-2004

C:Accession: E87269

R:Nieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kjolton, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.; Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of *Caulobacter crescentus*.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: E87269

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-267 <SNO>

A;Cross-references: UNIPROT:Q9ABR0; UNIPARC:UPI000000C6F59; GB:AE005673; NID:gl3421281; F

C;Genetics:

A;Gene: CC0166

C;Superfamily: sugar transferase, Exoy type

Query Match 70.8%; Score 17; DB 2; Length 267;
Best Local Similarity 25.0%; Pred. No. 1.8e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
20 CTTSSDC 27

RESULT 28

F70771

probable glutamateracemase - *Mycobacterium tuberculosis* (strain H37Rv)

C;Species: *Mycobacterium tuberculosis*

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C;Accession: F70771

R;Cole, S.T.; Broech, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: F70771

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-271 <COL>

A;Cross-references: UNIPROT:Q10626; UNIPARC:UPI000012FA0E; GB:Z73902; GB:AL123456; NID:9

A;Experimental source: strain H37Rv

C;Genetics:

A;Gene: murI

C;Superfamily: glutamate racemase

Query Match 70.8%; Score 17; DB 2; Length 271;
Best Local Similarity 25.0%; Pred. No. 1.8e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
75 CNSASAC 82

RESULT 29

S72790

glutamate racemase (EC 5.1.1.3) - *Mycobacterium leprae*

N;Alternate names: B1549_C2_210 protein

C;Species: *Mycobacterium leprae*

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C;Accession: S72790

R;Smith, D.R.; Robinson, K. submitted to the EMBL Data Library, November 1993

A;Description: *Mycobacterium leprae* cosmid B1549.

A;Reference number: S72582

A;Accession: S72790

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-272 <SMT>

A;Cross-references: UNIPROT:P46705; UNIPARC:UPI000012FA0D; EMBL:U00014; NID:g466903; PID

C;Genetics:

A;Gene: murI

C;Superfamily: glutamate racemase

C;Keywords: isomerase

Query Match 70.8%; Score 17; DB 1; Length 272;
Best Local Similarity 25.0%; Pred. No. 1.8e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
75 CNTASAC 82

RESULT 30

151172

transcription factor Rcc/EBP-2 - bullfrog

C;Species: *Rana catesbeiana* (bullfrog)

C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004

C;Accession: 151172

R;Chen, Y.; Hu, H.; Atkinson, B.G. Dev. Genet. 15, 366-377, 1994

A;Title: Characterization and expression of C/EBP-like genes in the liver of *Rana catesbeiana*

A;Reference number: 151171; MUID:95008407; PMID:7923939

A;Accession: 151172

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-284 <CHE>

A;Cross-references: UNIPROT:Q91295; UNIPARC:UPI00000FBAAC; EMBL:U08605; NID:g478890; PID

C;Superfamily: CCAAT/enhancer-binding protein alpha

Query Match 70.8%; Score 17; DB 2; Length 284;
Best Local Similarity 25.0%; Pred. No. 1.8e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
171 CTSNSAC 178

RESULT 31

C64795

citG protein - *Escherichia coli* (strain K-12)

C;Species: *Escherichia coli*

C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004

C;Accession: C64795

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Coj

A.; Rose, D.J.; Mau, B.; Shao, Y. Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of *Escherichia coli* K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: C64795

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-292 <BLAT>

A;Cross-references: UNIPROT:P77231; UNIPARC:UPI0000127A40; GB:AE000166; GB:U00096; NID:91

A;Experimental source: strain K-12, substrain MG1655

C;Genetics:

A;Gene: citG

C;Superfamily: citG protein

Query Match 70.8%; Score 17; DB 2; Length 292;
Best Local Similarity 25.0%; Pred. No. 1.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
142 CSTASFC 149

RESULT 32

D90710

2-(5'-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase (EC 4.2.-.-) *Escherichia coli*

C;Species: *Escherichia coli*

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C:Accession: D90710
 R:Hayashi, T.; Makino, K.; Onishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Gasawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shingana, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
 A:Reference number: A9629; MUID:21156231; PMID:11258796
 A:Accession: D90710
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-292 <HAY>
 A:Cross-references: UNIPROT:P58161; UNIPARC:UP10000127A3F; GB:BA000007; PIDN:BA034075.1;
 A:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: BC0652
 C:Superfamily: citG protein
 C:Keywords: carbon-oxygen lyase

Query Match 70.8%; Score 17; DB 2; Length 292;
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
 DB 142 CSTASFC 149

RESULT 33
 H85560
 2-(5'-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase (EC 4.2.-.-) citG - *Escherichia*
 C:Species: *Escherichia coli*
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 A:Accession: H85560
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grodzick, E.J.; Davis, N.W.; Lam, A.; Diallanita, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: H85560
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-292 <STO>
 A:Cross-references: UNIPROT:P58161; UNIPARC:UP10000127A3F; GB:AE005174; NID:G12513512; F
 A:Experimental source: strain O157:H7, substrain EDJ933
 C:Genetics:
 A:Gene: citG
 C:Superfamily: citG protein
 C:Keywords: carbon-oxygen lyase

Query Match 70.8%; Score 17; DB 2; Length 292;
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
 DB 142 CSTASFC 149

RESULT 34
 A98294
 hypothetical protein AGR_L_2619 [imported] - *Agrobacterium tumefaciens* (strain C58, Cere
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
 A:Accession: A98294
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: A98294
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-297 <KUN>
 A:Cross-references: UNIPROT:Q8UA53; UNIPARC:UP100000D222C; GB:AE007870; PIDN:AAK89875.1;

C:Genetics:
 A:Gene: AGR_L_2619
 A:Map position: linear chromosome

Query Match 70.8%; Score 17; DB 2; Length 297;
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
 DB 151 CAAVAAC 158

RESULT 35
 AH2989
 permease [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 A:Accession: AH2989
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AH2989
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-297 <KUN>
 A:Cross-references: UNIPROT:Q8UA53; UNIPARC:UP100000D222C; GB:AE006869; PIDN:AAL44334.1;
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu3521
 A:Map position: linear chromosome

Query Match 70.8%; Score 17; DB 2; Length 297;
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
 DB 151 CAAVAAC 158

RESULT 36
 AE0857
 conserved hypothetical protein STY3065 [imported] - *Salmonella enterica* subsp. *enterica*
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh
 A:Note: this species has also been called *Salmonella typhi*
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 A:Accession: AE0857
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 ; S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov.
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AE0857
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-305 <PAR>
 A:Cross-references: UNIPARC:UP1000005A39C; GB:AL513382; PIDN:CAD06044.1; PID:G16504011;
 C:Genetics:
 A:Gene: STY3065
 C:Superfamily: *Escherichia coli* hypothetical protein b2755

Query Match 70.8%; Score 17; DB 2; Length 305;
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
Db 179 CISAATSC 186

RESULT 37
hypotheoretical protein b2755 - Escherichia coli (strain K-12)

C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: G65056
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: G65056
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-305 <BLAT>
A:Cross-references: UNIPROT:Q46896; UNIPARC:UPI000013B034; GB:AE000359; GB:U00096; NID:G
A:Experimental source: strain K-12, substrain MG1655
C:Superfamily: Escherichia coli hypotheoretical protein b2755

Query Match 70.8%; Score 17; DB 2; Length 305;
Best Local Similarity 25.0%; Pred. No. 1.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
Db 179 CISAATSC 186

RESULT 38
B53925
hypotheoretical protein Z4064 [imported] - Escherichia coli (strain O157:H7, substrain EDL9

C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: B85925
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glanier, J.D.; Rose, D.J.; Mayhew
11ler, L.; Grobeck, B.J.; Davis, N.W.; Lim, A.; Diallanca, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: B85925
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-307 <STO>
A:Cross-references: UNIPROT:Q8X7W3; UNIPARC:UPI00000D09E4; GB:AE005174; NID:g12517212; F
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z4064

C:Superfamily: Escherichia coli hypotheoretical protein b2755

Query Match 70.8%; Score 17; DB 2; Length 307;
Best Local Similarity 25.0%; Pred. No. 1.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
Db 179 CISAATSC 186

RESULT 39
A91080
hypotheoretical protein ECs3609 [imported] - Escherichia coli (strain O157:H7, substrain RI

C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: A91080
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gend

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: A91080
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-307 <HAV>
A:Cross-references: UNIPROT:Q8X7W3; UNIPARC:UPI00000D09E4; GB:BA000007; PIDN:BA937032.1;
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs3609

C:Superfamily: Escherichia coli hypotheoretical protein b2755

Query Match 70.8%; Score 17; DB 2; Length 307;
Best Local Similarity 25.0%; Pred. No. 1.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
Db 179 CISAATSC 186

RESULT 40
T32376

hypotheoretical protein K10F12.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T32376
R:Wohlmann, P.; Beck, C.
submitted to the EMBL Data Library, September 1997
A:Description: The sequence of C. elegans coelmid K10F12.
A:Reference number: Z21157
A:Accession: T32376
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-309 <MOH>
A:Cross-references: UNIPROT:O17234; UNIPARC:UPI000007F612; EMBL:AF025462; PIDN:AAB71002.1
A:Experimental source: strain Bristol N2; clone K10F12
C:Genetics:
A:Gene: CESP:K10F12.4
A:Map position: 3
A:Insertions: 31/3; 123/2; 196/3; 239/1

Query Match 70.8%; Score 17; DB 2; Length 309;
Best Local Similarity 25.0%; Pred. No. 1.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
Db 7 CTSALUC 14

RESULT 41
B53522
20k cyclophilin - Toxoplasma gondii (fragment)

C:Species: Toxoplasma gondii
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: B53522
R:High, K.P.; Joiner, K.A.; Handschumacher, R.E.
J. Biol. Chem. 269, 9105-9112, 1994
A:Title: Isolation, cDNA sequences, and biochemical characterization of the major cyclophi
A:Reference number: A53522; MUID:94179329; PMID:8132648
A:Accession: B53522
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-347 <HTG>
A:Cross-references: UNIPROT:Q26995; UNIPARC:UPI000007DB6A; GB:U04634; NID:g436958; PID:g'
F:178-346/Domain: cyclophilin homology <CTP>

Query Match 70.8%; Score 17; DB 2; Length 347;
Best Local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8

Db 48 CASAEAC 55

RESULT 42

probable iron-siderophore uptake system transmembrane component - Streptomyces coelicolor

C/Species: Streptomyces coelicolor

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C/Accession: T36890

R/Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, August 1999

A/Reference number: 221617

A/Accession: T36890

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-348 <MUR>

A/Cross-references: UNIPROT:O9S213; UNIPARC:UPI00000DB327; EMBL:AL109848; P1DN:CA52851.

A/Experimental source: strain A3(2)

C/Genetics:

A/Genetic: SCOEDB:SC151.27C

C/Superfamily: vitamin B12 transport protein btuc

Query Match

Best Local Similarity 70.8%; Score 17; DB 2; Length 348;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 1 CXXXXXXC 8

262 CGAATAC 269

RESULT 43

hypothetical protein At2g17120 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C/Accession: C84548

R/Li, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.

ause, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: C84548

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-350 <STO>

A/Cross-references: UNIPROT:O23006; UNIPARC:UPI00000A22DD; GB:AE002093; NID:g2558660; PI

A/Genetics:

A/Genetic: At2g17120

A/Map position: 2

Query Match

Best Local Similarity 70.8%; Score 17; DB 2; Length 350;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 1 CXXXXXXC 8

31 CGSSTSTC 38

RESULT 44

conserved hypothetical protein CC2083 [imported] - Caulobacter crescentus

C/Species: Caulobacter crescentus

C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C/Accession: B87507

R/Nierman, W.C.; Feldlyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolor

n, U.; Emlaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A/Title: Complete Genome Sequence of Caulobacter crescentus.

A/Reference number: A87249; MUID:21173698; PMID:11259647

A/Accession: B87507

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-352 <STO>

A/Cross-references: UNIPROT:Q9A6L0; UNIPARC:UPI00000C75D3; GB:AE005673; NID:g13423564; P

A/Genetics:

A/Genetic: CC2083

Query Match

Best Local Similarity 70.8%; Score 17; DB 2; Length 352;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 1 CXXXXXXC 8

11 CAALISAC 18

RESULT 45

conserved hypothetical protein TP0730 - syphilis spirochete

C/Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C/Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C/Accession: G71287

R/Praser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin

ison, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo

they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998

A/Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A/Reference number: A71250; MUID:9832770; PMID:9665876

A/Accession: G71287

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-365 <COL>

A/Cross-references: UNIPROT:O83712; UNIPARC:UPI00000C0AA2; GB:AE001245; GB:AE000520; NID

A/Experimental source: strain Nichols

C/Genetics:

A/Genetic: TP0730

Query Match

Best Local Similarity 70.8%; Score 17; DB 2; Length 365;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 1 CXXXXXXC 8

162 CAASDC 169

RESULT 46

nifs family enzyme (cysteine desulfurase/cysteine sulfinate desulfinase) [imported] - Cl

nifs

C/Species: Clostridium acetobutylicum

C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004

C/Accession: D97175

R/Nolling, U.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J. Daily, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A/Reference number: A96900; MUID:21359325; PMID:21359325

A/Accession: D97175

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-376 <KOR>

A/Cross-references: UNIPROT:O97GY1; UNIPARC:UPI00000D7546; GB:AE001437; P1DN:AAK80191.1;

A/Experimental source: Clostridium acetobutylicum ATCC824

C/Genetics:

A/Genetic: CAC2234

C/Superfamily: nitrogen fixation protein nifs

Query Match

Best Local Similarity 70.8%; Score 17; DB 2; Length 376;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 1 CXXXXXXC 8

Db 319 CASGSAC 326

RESULT 47

C34443

nitrogenase cofactor synthesis protein nifs - Anabaena sp.

N:Contains: L-cysteine sulfotransferase (EC 2.8.1.-)

C:Species: Anabaena sp.

C:Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 09-Jul-2004

C:Accession: C34443; B32361

R:Muligan, M.E.; Haselkorn, R.

J. Biol. Chem. 264, 19200-19207, 1989

A:Title: Nitrogen fixation (nif) genes of the cyanobacterium Anabaena species strain PCC

A:Reference number: A34443; MUID:90037054; PMID:2553733

A:Accession: C34443

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-400 <MDL>

A:Cross-references: UNIPROT:P12623; UNIPARC:UPI000016555C; GB:J05111; NID:G142034; PIDN:

R:Muligan, M.E.; Bulkema, W.J.; Haselkorn, R.

J. Bacteriol. 170, 4406-4410, 1988

A:Title: Bacterial-type ferredoxin genes in the nitrogen fixation regions of the cyanoba

A:Reference number: A31884; MUID:88314954; PMID:2842320

A:Accession: B32361

A:Molecule type: DNA

A:Residues: 1-7 <MD2>

A:Cross-references: UNIPARC:UPI0000178A6F

C:Superfamily: nitrogen fixation protein nifs

C:Keywords: phosphoprotein; pyridoxal phosphate (lys) (covalent) #status predicted

F:324/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

F:324/Active site: Cys (cysteine persulfide intermediate) #status predicted

Query Match

Best Local Similarity 70.8%; Score 17; DB 2; Length 400;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8

Db 317 CASGSAC 324

RESULT 48

AE1988

nitrogenase cofactor synthesis protein (imported) - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C:Accession: AE1988

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Iriguchi,

Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AE1988

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-400 <KUR>

A:Cross-references: UNIPROT:P12623; UNIPARC:UPI000013010C; GB:BA000019; PIDN:BA073413.1;

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: nifs

C:Superfamily: nitrogen fixation protein nifs

Query Match 70.8%; Score 17; DB 2; Length 400;

Best Local Similarity 25.0%; Pred. No. 2e+02; Mismatches 6; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8

Db 317 CASGSAC 324

RESULT 49

WNVIA

immediate-early protein IE-N - Autographa californica nuclear polyhedrosis virus

C:Species: Autographa californica nuclear polyhedrosis virus, AcNPV

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004

C:Accession: A39150

R:Carson, D.D.; Summers, M.D.; Guarino, L.A.

Virology 182, 279-286, 1991

A:Title: Molecular analysis of a baculovirus regulatory gene.

A:Reference number: A39150; MUID:91220660; PMID:2024466

A:Accession: A39150

A:Molecule type: DNA

A:Residues: 1-408 <CAR>

A:Cross-references: UNIPROT:P24647; UNIPARC:UPI0000170DEB; GB:M59422; NID:G332437; PIDN:

C:Superfamily: AcNPV immediate-early protein IE-N

C:Keywords: DNA binding; early protein; tandem repeat; transcription regulation

F:34-49/Region: 7-residue repeats

F:51-58/Region: 4-residue repeats

F:190-196/Region: glutamine-rich

Query Match 70.8%; Score 17; DB 1; Length 408;

Best Local Similarity 25.0%; Pred. No. 2.1e+02; Mismatches 6; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8

Db 253 CSATSDC 260

RESULT 50

A72869

early gene transactivator - Autographa californica nuclear polyhedrosis virus

C:Species: Autographa californica nuclear polyhedrosis virus, AcNPV

A:Note: dsDNA virus

C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 09-Jul-2004

C:Accession: A72869

R:AYres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.

Virology 202, 586-605, 1994

A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.

A:Reference number: A72869; MUID:94303173; PMID:8030224

A:Accession: A72869

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-408 <AYR>

A:Cross-references: UNIPROT:P24647; UNIPARC:UPI00001387E9; GB:L22858; NID:G510708; PIDN:

C:Genetics:

A:Gene: Ac-IE-2

C:Superfamily: AcNPV immediate-early protein IE-N

Query Match 70.8%; Score 17; DB 2; Length 408;

Best Local Similarity 25.0%; Pred. No. 2.1e+02; Mismatches 6; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8

Db 253 CSATSDC 260

Search completed: January 4, 2006, 16:10:26

Job time : 9.27826 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 / Search time 52.2435 seconds
(without alignments)
108.037 Million cell updates/sec

Title: US-09-932-322-9
Perfect score: 24
Sequence: 1 CXXXXXXC 8

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database: UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	75.0	69	Q9N669_GCAEN	Q9N669 conus ebrae
2	18	75.0	69	Q9NCS8_GCAEN	Q9NCS8 conus ebrae
3	18	75.0	69	Q9NCS9_GCAEN	Q9NCS9 conus ebrae
4	18	75.0	69	Q9NCT2_GCAEN	Q9NCT2 conus ebrae
5	18	75.0	69	Q9NCT3_GCAEN	Q9NCT3 conus ebrae
6	18	75.0	69	Q9VD36_DROME	Q9VD36 anopheles g
7	18	75.0	90	Q7QFP9_ANOGA	Q7QFP9 anopheles g
8	18	75.0	90	Q9B111_ANOGA	Q9B111 anopheles g
9	18	75.0	114	Q82854_PSETR	Q82854 jembrana di
10	18	75.0	161	Q9EQJ8_MOUSE	Q9EQJ8 mus musculu
11	18	75.0	163	Q7QPM4_GIALA	Q7QPM4 giardia lam
12	18	75.0	177	Q8LDT9_ARATH	Q8LDT9 arabidopsis
13	18	75.0	177	Q9C9L7_ARATH	Q9C9L7 arabidopsis
14	18	75.0	199	Q6OM47_CABBR	Q6OM47 caenorhabdi
15	18	75.0	200	Q5VMP6_ORYSA	Q5VMP6 oryza sativ
16	18	75.0	224	Q9EQJ7_MOUSE	Q9EQJ7 mus musculu
17	18	75.0	233	Q6KZG5_ORYSA	Q6KZG5 oryza sativ
18	18	75.0	242	Q51XD3_MAGGR	Q51XD3 magnaporthe
19	18	75.0	245	Q6BCP3_HPBVO	Q6BCP3 hepatictis b
20	18	75.0	258	Q9Y0E9_DROME	Q9Y0E9 drosophila
21	18	75.0	269	Q7PR07_ANOGA	Q7PR07 anopheles g
22	18	75.0	319	Q7Y198_ORYSA	Q7Y198 oryza sativ
23	18	75.0	320	Q6L1A5_CABBR	Q6L1A5 caenorhabdi
24	18	75.0	328	Q6ZNM3_HUMAN	Q6ZNM3 homo sapien
25	18	75.0	332	Q6Z4B6_ORYSA	Q6Z4B6 oryza sativ
26	18	75.0	350	Q94589_9CILI	Q94589 lambdaion b
27	18	75.0	355	RHOM_DROME	P20350 drosophila
28	18	75.0	355	Q540V7_DROME	Q540V7 drosophila
29	18	75.0	367	TTSD_MOUSE	P23949 mus musculu
30	18	75.0	404	Q86SB6_DROVI	Q86SB6 drosophila
31	18	75.0	417	Q6Z4E5_ORYSA	Q6Z4E5 oryza sativ

32	18	75.0	445	Q8UTL4_ORYSA	Q8UTL4 oryzias lat
33	18	75.0	480	Q7SHC6_NEUTR	Q7SHC6 neurospora
34	18	75.0	480	Q758Y3_ASHCO	Q758Y3 ashbya goss
35	18	75.0	492	TTSD_HUMAN	P47974 homo sapien
36	18	75.0	494	Q53TE4_HUMAN	Q53TE4 homo sapien
37	18	75.0	526	Q6PD21_MOUSE	Q6PD21 mus musculu
38	18	75.0	544	Q5B195_DROME	Q5B195 drosophila
39	18	75.0	545	Q51TR2_TUHV1	Q51TR2 tupaiid her
40	18	75.0	551	Q5Z8N9_ORYSA	Q5Z8N9 oryza sativ
41	18	75.0	551	Q8SB30_ORYSA	Q8SB30 oryza sativ
42	18	75.0	552	Q5AP87_CANAL	Q5AP87 candida alb
43	18	75.0	560	Q8UJ53_HPBVO	Q8UJ53 hepatictis b
44	18	75.0	581	Q6N1S1_CORDI	Q6N1S1 corynebacte
45	18	75.0	592	Q51VZ0_MAGGR	Q51VZ0 magnaporthe
46	18	75.0	606	Q6Z4E7_ORYSA	Q6Z4E7 oryza sativ
47	18	75.0	691	Q4RNV9_TETNG	Q4RNV9 tetradon n
48	18	75.0	705	Q7S136_NEUTR	Q7S136 neurospora
49	18	75.0	711	Q5RFW2_BRARE	Q5RFW2 brachydanio
50	18	75.0	722	Q5AP84_CANAL	Q5AP84 candida alb
51	18	75.0	743	Q09517_HPBVO	Q09517 hepatictis b
52	18	75.0	769	Q80GK5_HPBVO	Q80GK5 hepatictis b
53	18	75.0	787	Q09511_HPBVO	Q09511 hepatictis b
54	18	75.0	795	Q8V1H4_HPBVO	Q8V1H4 hepatictis b
55	18	75.0	799	Q4T7H8_TETNG	Q4T7H8 tetradon n
56	18	75.0	801	Q09504_HPBVO	Q09504 hepatictis b
57	18	75.0	801	Q09505_HPBVO	Q09505 hepatictis b
58	18	75.0	823	Q09QK3_HPBVO	Q09QK3 hepatictis b
59	18	75.0	836	Q8V1I3_HPBVO	Q8V1I3 hepatictis b
60	18	75.0	837	Q50UT2_HPBVO	Q50UT2 hepatictis b
61	18	75.0	843	DPOL_HPBVO	P17394 hepatictis b
62	18	75.0	843	DPOL_HPBVO	P17395 hepatictis b
63	18	75.0	843	Q09509_HPBVO	Q09509 hepatictis b
64	18	75.0	843	Q39877_HPBVO	Q39877 hepatictis b
65	18	75.0	843	Q39882_HPBVO	Q39882 hepatictis b
66	18	75.0	843	Q91527_HPBVO	Q91527 hepatictis b
67	18	75.0	843	Q5DW06_HPBVO	Q5DW06 hepatictis b
68	18	75.0	843	Q5DW09_HPBVO	Q5DW09 hepatictis b
69	18	75.0	843	Q5DM13_HPBVO	Q5DM13 hepatictis b
70	18	75.0	843	Q5Q0T5_HPBVO	Q5Q0T5 hepatictis b
71	18	75.0	843	Q67937_HPBVO	Q67937 hepatictis b
72	18	75.0	843	Q68R01_HPBVO	Q68R01 hepatictis b
73	18	75.0	843	Q68R02_HPBVO	Q68R02 hepatictis b
74	18	75.0	843	Q68R03_HPBVO	Q68R03 hepatictis b
75	18	75.0	843	Q68R04_HPBVO	Q68R04 hepatictis b
76	18	75.0	843	Q68R09_HPBVO	Q68R09 hepatictis b
77	18	75.0	843	Q68R03_HPBVO	Q68R03 hepatictis b
78	18	75.0	843	Q762E4_HPBVO	Q762E4 hepatictis b
79	18	75.0	843	Q765V6_HPBVO	Q765V6 hepatictis b
80	18	75.0	843	Q76B37_HPBVO	Q76B37 hepatictis b
81	18	75.0	843	Q77DP5_HPBVO	Q77DP5 hepatictis b
82	18	75.0	843	Q77DP8_HPBVO	Q77DP8 hepatictis b
83	18	75.0	843	Q77D01_HPBVO	Q77D01 hepatictis b
84	18	75.0	843	Q77D07_HPBVO	Q77D07 hepatictis b
85	18	75.0	843	Q77D80_HPBVO	Q77D80 hepatictis b
86	18	75.0	843	Q77D81_HPBVO	Q77D81 hepatictis b
87	18	75.0	843	Q77D81_HPBVO	Q77D81 hepatictis b
88	18	75.0	843	Q77D81_HPBVO	Q77D81 hepatictis b
89	18	75.0	843	Q77HP5_HPBVO	Q77HP5 hepatictis b
90	18	75.0	843	Q77H00_HPBVO	Q77H00 hepatictis b
91	18	75.0	843	Q77H00_HPBVO	Q77H00 hepatictis b
92	18	75.0	843	Q805G6_HPBVO	Q805G6 hepatictis b
93	18	75.0	843	Q80G08_HPBVO	Q80G08 hepatictis b
94	18	75.0	843	Q80H00_HPBVO	Q80H00 hepatictis b
95	18	75.0	843	Q80H20_HPBVO	Q80H20 hepatictis b
96	18	75.0	843	Q80H48_HPBVO	Q80H48 hepatictis b
97	18	75.0	843	Q80U60_HPBVO	Q80U60 hepatictis b
98	18	75.0	843	Q80J72_HPBVO	Q80J72 hepatictis b
99	18	75.0	843	Q80J78_HPBVO	Q80J78 hepatictis b
100	18	75.0	843	Q80J80_HPBVO	Q80J80 hepatictis b
101	18	75.0	843	Q80J83_HPBVO	Q80J83 hepatictis b
102	18	75.0	843	Q80M08_HPBVO	Q80M08 hepatictis b
103	18	75.0	843	Q81116_HPBVO	Q81116 hepatictis b
104	18	75.0	843	Q81120_HPBVO	Q81120 hepatictis b


```
981 16 66.7 122 2 063fj9_BAC2Z 063fj9 bacillus ce
982 16 66.7 122 2 05uvz1_9ADEN 05uvz1 human adeno
983 16 66.7 122 2 077923_ADB35 077923 human adeno
984 16 66.7 122 2 08B850_9ADEN 08B850 human adeno
985 16 66.7 123 1 11z21_PBRMA 080xg2 petromyscus
986 16 66.7 123 2 07Pp38_ANOGA 07Pp38 apophies g
987 16 66.7 123 2 08JG94_GINCI 08JG94 ginglymosto
988 16 66.7 124 2 08PVO6_METMA 08PVO6 methamosarc
989 16 66.7 124 2 07R266_GIALA 07R266 giardia lam
990 16 66.7 124 2 016121_TENMO 016121 tenebrio mo
991 16 66.7 124 2 062710_CABBR 062710 caenorhabdi
992 16 66.7 124 2 09U744_TENMO 09U744 tenebrio mo
993 16 66.7 125 2 06DLX5_TENMO 06DLX5 tenebrio mo
994 16 66.7 125 2 0814J7_CABEL 0814J7 caenorhabdi
995 16 66.7 125 2 038346_BPLHL 038346 lactococcus
996 16 66.7 125 2 04NR72_9BELT 04NR72 anaeromyxob
997 16 66.7 125 2 08JFJ7_GINCI 08JFJ7 ginglymosto
998 16 66.7 125 2 08JFJ8_GINCI 08JFJ8 ginglymosto
999 16 66.7 125 2 08JFJ8_GINCI 08JFJ8 ginglymosto
1000 16 66.7 125 2 08JFJ8_GINCI 08JFJ8 ginglymosto
```

ALIGNMENTS

```
RESULT 1
Q9N669_9CAEN PRELIMINARY; PRT; 69 AA.
```

```
AC 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE Four-loop conotoxin (Fragment).
OS Conus ebraeus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=89425;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=20416305; PubMed=10958845;
RA Duda T.F., Jr., Palumbi S.R.;
RT "Evolutionary diversification of multigene families: allelic selection
of toxins in predatory cone snails."
RL Mol. Biol. Evol. 17:1286-1293(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Duda T.F., Palumbi S.R.;
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF174281; AAF89945.1; -; mRNA.
DR EMBL; AF174280; AAF89944.1; -; mRNA.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0008200; F:ion channel inhibitor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR004214; Conotoxin.
DR Pfam; PF02950; Conotoxin; 1.
FT NON TER
SQ SEQUENCE 69 AA; 7662 MW; C8C826FDC920C4F1 CRC64;
```

```
Query Match 75.0%; Score 18; DB 2; Length 69;
Best Local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 1 CXXXXXXC 8
Db 61 CSTATSTC 68
```

```
RESULT 2
Q9NCS8_9CAEN PRELIMINARY; PRT; 69 AA.
AC Q9NCS8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
```

```
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Four-loop conotoxin (Fragment).
OS Conus ebraeus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=89425;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=20416305; PubMed=10958845;
RA Duda T.F., Jr., Palumbi S.R.;
RT "Evolutionary diversification of multigene families: allelic selection
of toxins in predatory cone snails."
RL Mol. Biol. Evol. 17:1286-1293(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Duda T.F., Palumbi S.R.;
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF174287; AAF89951.1; -; mRNA.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0008200; F:ion channel inhibitor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR004214; Conotoxin.
DR Pfam; PF02950; Conotoxin; 1.
FT NON TER
SQ SEQUENCE 69 AA; 7674 MW; 83D526FDC934422A CRC64;
```

```
Query Match 75.0%; Score 18; DB 2; Length 69;
Best Local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 1 CXXXXXXC 8
Db 61 CSTATSTC 68
```

```
RESULT 3
Q9NCS9_9CAEN PRELIMINARY; PRT; 69 AA.
```

```
AC 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Four-loop conotoxin (Fragment).
OS Conus ebraeus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=89425;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=20416305; PubMed=10958845;
RA Duda T.F., Jr., Palumbi S.R.;
RT "Evolutionary diversification of multigene families: allelic selection
of toxins in predatory cone snails."
RL Mol. Biol. Evol. 17:1286-1293(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Duda T.F., Palumbi S.R.;
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF174286; AAF89950.1; -; mRNA.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0008200; F:ion channel inhibitor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR004214; Conotoxin.
DR Pfam; PF02950; Conotoxin; 1.
FT NON TER
SQ SEQUENCE 69 AA; 7650 MW; DF3826FDC920C4F1 CRC64;
```

```
Query Match 75.0%; Score 18; DB 2; Length 69;
Best Local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy      1 CXXXXXXC 8
Db      61 CSTATSTC 68
```

RESULT 4	
Q9NCT2_9CAEN	
ID Q9NCT2_9CAEN PRELIMINARY;	
PRT;	69 AA

DT 01-OCT-2000 (TEMBBprel. 15, Created)
DT 01-OCT-2000 (TEMBBprel. 15, last sequence update)
DT 01-OCT-2003 (TEMBBprel. 25, last annotation update)
DE Four-loop conotoxin (fragment).
OS Conus ebraeus.
OC Eukaryote; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeconcha; Hypsogastropoda;
OC Neogastropoda; Comidae; Conidae; Conus.
NCBI_TaxID=89425;

RESULT 5		99NCNT3_9CAEN	
ID	99NCNT3_9CAEN	PRELIMINARY;	PRT; 69 AA.
AC	99NCNT3;		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)		
DE	Four-loop conotoxin (Fragment).		
OS	Conus ebraeus.		
OS	Eukaryote; Metazoa; Mollusca; Gastropoda; Orthogastropoda;		
OC	Agastropoda; Caenogastropoda; Sorbeoconcha; Hypogastropoda;		
OC	Neogastropoda; Conoidea; Conidae; Conus.		
OX	NCBI_TaxID=89425;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=20416305; PubMed=10958845;		
RA	Duda T.F., Ufr., Palumbi S.R.;		
RT	"Evolutionary diversification of multigene families: allelic selection		
RT	of toxins in predatory cone snails.";		
RL	Mol. Biol. Evol. 17:1286-1293(2000).		
RP	[2]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Duda T.F., Palumbi S.R.;		
RL	Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.		
DR	EMBL; AF174279; AAFF89943.1; -; mRNA.		

DR	GO; GO:0005276; C:extracellular region; IEA.
DR	GO; GO:0008505; F:ion channel inhibitor activity; IEA.
DR	GO; GO:0009405; P:phthogenesis; IEA.
DR	InterPro: IPR004214; Conotoxin.
DR	Pfam: PF02950; Conotoxin; 1.
FT	NON_TER
FO	SEQUENCE 69 AA; 7643 MW; D5699CEDC920CAFI CRC64;

Q9V336	DROME		
1D	Q9V336_DROME	PRELIMINARY;	PRT; 69 AA.
AC	Q9V336		
DT	01-MAY-2000	(TrEMBLrel. 13, Created)	
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)	
DT	01-MAY-2000	(TrEMBLrel. 13, Last annotation update)	
DE	CG13858-PA.		
GN	Name=CG13858; ORFNames=CG13858;		
OS	Drosophila melanogaster (fruit fly)		
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	NCBLOTTIDE SQUENCE.		
RX	MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;		
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Miklos B.D.,		
RA	Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., McKis G.L.G.,		
RA	Abvill J.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,		
RA	Ballew R.M., Baau A.A., Bakendell J., Bayraktaroglu I., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brinkstein P., Brotler P.,		
RA	Buttis K.C., Busam D.A., Butler H., Cadelu E., Center A., Chandra I.,		
RA	Cherry J.M., Cusale S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pallos B., Delcher A., Deng Z., Mayh A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,		
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA	Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,		
RA	Holtin D., Houston K.A., Howland T.J., Wei M.-H., Ilegwam C.,		
RA	Jacobs M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C.J., Kravitz S., Kulp D., Lai Z.,		
RA	Iakou P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Milshina N.V., Mobery C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,		
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reinert K., Remington K., Saunders R.D.C., Scheele F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spiet E., Spreading A.C., Stapleton M., Strong R., Sun B.,		
RA	Svirskas R., Tector C., Turner K., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,		
RA	Williams S.M., Wozniak T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;		
RT	"The genome sequence of <i>Drosophila melanogaster</i> .";		
RL	Science 287:2185-2195(2000).		
RN	[2]		

AC Q82854;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN Name=tat;
OS Jembrana disease virus.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Bovine lentiviruses.
OC NCBI_TaxID=36370;
OK [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Tabanan/87;
RX MEDLINE=97201584; PubMed=9049370;
RA Chadwick B.J., Coelen R.J., Wilcox G.E., Samuels L.M., Kertyadnya G.,
RT "Nucleotide sequence analysis of Jembrana disease virus: a bovine
RT lentivirus associated with an acute disease syndrome.";
RL J. Gen. Virol. 76:1637-1650(1995).
DR EMBL; U21603; AAA64395.1; -, Genomic RNA.
DR GO; GO:0042025; C:host cell nucleus; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR001831; IV_Tat.
DR Pfam; PF00539; Tat; 1.
KM Activator; Nuclear protein; RNA-binding; Transcription;
KM Transcription regulation.
SQ SEQUENCE 114 AA; 12457 MW; 77A7CB6FAF128D5A CRC64;

Query Match 75.0%; Score 18; DB 2; Length 114;
Best Local Similarity 25.0%; Pred. No. 2.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 96 CASSASC 103

RESULT 10
Q9EQJ8 MOUSE
ID Q9EQJ8 MOUSE PRELIMINARY; PRT; 161 AA.
AC Q9EQJ8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE T1S11D deletion variant (Fragment).
GN Name=zfp3612; Synonyms=Btf2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
OK [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BLKS/J;
RA Cho K., Hobson K., Greenhalgh D.G.;
RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF205222; AAG45250.1; -, mRNA.
DR HSSP; P22893; IMGO.
DR SMR; Q9EQJ8; 1-49.
DR MGI; MGI:107945; Zfp3612.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR000571; Znf_CCH.
DR Pfam; PF00642; zf-CCH; 1.
DR SMART; SMO0356; Znf_C3H1; 1.
FT NON TER 1 1
KM NON TER 161 161
SQ SEQUENCE 161 AA; 17101 MW; BF09DF73709215EE CRC64;

Query Match 75.0%; Score 18; DB 2; Length 161;
Best Local Similarity 25.0%; Pred. No. 2.8e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CXXXXXXC 8
Db 129 CASSASC 136

RESULT 11
Q7QPM4 GIALA
ID Q7QPM4 GIALA PRELIMINARY; PRT; 163 AA.
AC Q7QPM4;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GLP 54 16521 16030.
OS Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
OC NCBI_TaxID=184922;
OK [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=WB C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RA Olsen G.J., Sogin M.L.;
RT "Draft sequence of the Giardia lamblia genome.";
RT Submitted (MAR-2003) to the EMBL/Genbank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAC801000166; EAA36969.1; -, Genomic DNA.
SQ SEQUENCE 163 AA; 17478 MW; 3AF4A5F8BD9A910C CRC64;

Query Match 75.0%; Score 18; DB 2; Length 163;
Best Local Similarity 25.0%; Pred. No. 2.8e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 24 CAATTAAC 31

RESULT 12
Q8LD79 ARATH
ID Q8LD79 ARATH PRELIMINARY; PRT; 177 AA.
AC Q8LD79;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein.
OS Arabidopsis thaliana (Mouse-ear cress);
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
OK [1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=22086475; PubMed=12093376;
RX Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.,
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RP Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY085804; AAM63020.1; -, mRNA.
KM Hypothetical protein.
SQ SEQUENCE 177 AA; 19819 MW; CCA7A7841CBAEBB38 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 177;
Best Local Similarity 25.0%; Pred. No. 2.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
SQ
Db 87 CSSSSSSC 94

RESULT 13

09C9L7 ARATH PRELIMINARY; PRT: 177 AA.
AC 09C9L7 ARATH PRELIMINARY; PRT: 177 AA.
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein T6C23.4.
GN Name=T6C23.4;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Uteirack T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Town C.D., Kaul S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC013289; AAG52545.1; -; Genomic_DNA.
DR PIR; F96719; F96719.
KW Hypothetical protein.
SQ SEQUENCE 177 AA; 19834 MW; 03D870608FEB8246 CRC64;
Query Match 75.0%; Score 18; DB 2; Length 177;
Best Local Similarity 25.0%; Pred. No. 2.9e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Oy 1 CXXXXXXC 8
Db 87 CSSSSSSC 94

RESULT 14

060M47 CABR PRELIMINARY; PRT: 199 AA.
AC 060M47;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Hypothetical protein CBG23307.
GN Name=CBG23307;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 BPTI/Kunitz inhibitor domains.
DR EMBL; CAAC010004.1; CAE75329.1; -; Genomic DNA.
DR GO; GO:0004867; F:serine-type endopeptidase inhibitor activity; IEA.
DR InterPro; IPR002223; Prot_inh_kunz-m.
DR InterPro; IPR006150; Worm_repeat_1.
DR Pfam; PF00014; Kunitz_BPTI; 2.
DR PRINTS; PR00759; BASICPRASE.
DR PRODOM; PD000222; Prot_inh_kunz-m; 2.
DR SMART; SM00131; KU; 2.
DR SMART; SM00289; WRI; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS50279; BPTI_KUNITZ_2; 2.

KW Hypothetical protein.
SQ SEQUENCE 199 AA; 21668 MW; 380A14B890A152A3 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 199;
Best Local Similarity 25.0%; Pred. No. 3.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
Db 112 CSSAATTC 119

RESULT 15

OSVMP6 ORYSA PRELIMINARY; PRT: 200 AA.
AC OSVMP6 ORYSA PRELIMINARY; PRT: 200 AA.
DT 01-FEB-2005 (TRENBLrel. 29, Created)
DT 01-FEB-2005 (TRENBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
DE Hypothetical protein OSJNB0008D07.38.
GN Name=OSJNB0008D07.38;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsunoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikono M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa M., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizudayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsunoto T., Katayose Y.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006237; BAD69279.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 200 AA; 23098 MW; 0E4249F62A3678AA CRC64;
Query Match 75.0%; Score 18; DB 2; Length 200;
Best Local Similarity 25.0%; Pred. No. 3.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Oy 1 CXXXXXXC 8
Db 102 CTAATTC 109

RESULT 16

09EOL7 MOUSE PRELIMINARY; PRT: 224 AA.
AC 09EOL7 MOUSE PRELIMINARY; PRT: 224 AA.
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE TIS1D insertion variant (Fragment).
GN Name=Tsp3612; Synonyms=Btf2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

```

OC Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BLKS/J;
RA Cho K., Hobson K., Greenhalgh D.G.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF205223; AAG45251.1; -; mRNA.
DR HSSP; P22893; 1M90.
DR SMR; Q9EQL7; 1-49.
DR MGI; MGI:107945; Zfp3612.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR000571; Znf_CCH.
DR Pfam; PF00642; Zf_CCH; 1.
DR SMART; SM00356; Znf_C3H1; 1.
FT NON_TER 1
FT SEQUENCE 224 AA; 23018 MW; 26E09C8465A5A61E CRC64;
SQ
Query Match 75.0%; Score 18; DB 2; Length 224;
Best Local Similarity 25.0%; Pred. No. 3.2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 129 CASSASC 136

RESULT 17
Q6KRG5 ORYSA PRELIMINARY; PRT; 233 AA.
ID Q6KRG5;
AC Q6KRG5;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Glucanase-1 like.
GN Name=OSUNBA0052M16.18;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriactoidae; Oryzae; Oryza.
OC NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
clone:OSUNBA0052M16."
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005841; BAD23651.1; -; Genomic_DNA.
DR Gramene; Q6KRG5; -; 24907 MW; E4805DB35FD0235 CRC64;
SQ
Query Match 75.0%; Score 18; DB 2; Length 233;
Best Local Similarity 25.0%; Pred. No. 3.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 199 CAATTAAC 206

RESULT 18
OS1XD3 MAGR PRELIMINARY; PRT; 242 AA.
ID OS1XD3;
AC OS1XD3;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Predicted protein.
ORNames=M607952.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.

```

```

OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armstrong J., Bachantang P., Baldwin J., Barry A.,
RA Bayul T., Bilshteyn B., Bloom T., Blye J., Boguslavsky L.,
RA Boroweky M., Bouhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshtang Y., Citroen M.,
RA Collamore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Geatin G., Gierre S.,
RA Gurke A., Guyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Huby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysellis M., Karlsson E.,
RA Kells C., Kieu A., Kistner P., Kodira C., Kulbokas E., Labutli K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Linblad-toh K., Liu X., Lokysang Y., Lokysang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrum J., Menus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelson T., Mlenga V., Moru K.,
RA Mozes J., Mullrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okawa O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunhkhong P., Pigani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipali S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougen C.,
RA Spencer B., Stalker J., Strange-thomson N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchunga P.,
RA Tenzing P., Testaye S., Theodore J., Thoultsang Y., Topham K.,
RA Towey S., Tsamla T., Teomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea."
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01000977; EAA53675.1; -; Genomic DNA.
DR SEQUENCE 242 AA; 25532 MW; 3393ABE816D10BA3 CRC64;
SQ
Query Match 75.0%; Score 18; DB 2; Length 242;
Best Local Similarity 25.0%; Pred. No. 3.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 62 CAASSSC 69

RESULT 19
Q6BCP3 HPBV0 PRELIMINARY; PRT; 245 AA.
ID Q6BCP3;
AC Q6BCP3;

```

DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Pre-S/S protein (Fragment).
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
NCBI_TaxID=10407;
GN [1]
RP NUCLEOTIDE SEQUENCE.
RA Nguyen H.C., Dong S.H., Ho T.T.T., Ho H.T.D.;
RT "Sequence analysis of pre-S/S gene segment from some clinical
RT hepatitis B virus isolates."
RL Submitted (JUN-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY665569; AAT77983.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR GO; GO:0016032; P:viral life cycle; IEA.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000349; Hepvir_surfa_g.
DR PANTHER; PTHR10832; Hepvir_surfa_g. 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER 1 1
FT NON_TER 245 245
SQ SEQUENCE 245 AA; 26093 MW; C275D837A352298E CRC64;

Query Match 75.0%; Score 18; DB 2; Length 245;
Best Local Similarity 25.0%; Pred. No. 3.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 80 CASSTSSC 87

RESULT 20
Q9Y0E9 DROME PRELIMINARY; PRT; 258 AA.
ID Q9Y0E9 DROME PRELIMINARY; PRT; 258 AA.
AC Q9Y0E9;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Salivary gland secretion protein (Fragment).
GN Name=Sgs1, Synonyms=Sgs-1, ORNames=CG3047;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
GN [1]
RP NUCLEOTIDE SEQUENCE.
RA Roch G.E., Wachtler S., Bornschein H., Lehmann M., Korge G.;
RT "Structure and regulation of the salivary gland secretion protein gene
RT Sgs-1 of Drosophila melanogaster."
RL Genetics 0:0-0(1999).
DR EMBL; AF156227; AAD43809.1; -; Genomic_DNA.
DR FlyBase; FBgn0003372; CG3047.
DR FlyBase; FBgn0003372; Sgs1.
FT NON_TER 258 258
SQ SEQUENCE 258 AA; 27832 MW; 0D9BF9218E6742EC CRC64;

Query Match 75.0%; Score 18; DB 2; Length 258;
Best Local Similarity 25.0%; Pred. No. 3.4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 226 CTTTSSC 233

RESULT 21
Q7PRQ7 ANOGA PRELIMINARY; PRT; 269 AA.
ID Q7PRQ7 ANOGA PRELIMINARY; PRT; 269 AA.
AC Q7PRQ7;
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE ENSANGP0000001657 (Fragment).
GN ORFNames=ENSANG0000001387;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
NCBI_TaxID=180454;
GN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (APR-2004) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.
CC [2]
CC The Anopheles gambiae Sequence Committee;
CC Submitted (APR-2004) to the EMBL/Genbank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB0100847; EAA06779.3; -; Genomic_DNA.
FT NON_TER 1 1
FT NON_TER 269 269
SQ SEQUENCE 269 AA; 24228 MW; 25BBFF71FD71F1F2 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 269;
Best Local Similarity 25.0%; Pred. No. 3.4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 55 CSSTSSC 62

RESULT 22
Q7Y198 ORYSA PRELIMINARY; PRT; 319 AA.
ID Q7Y198 ORYSA PRELIMINARY; PRT; 319 AA.
AC Q7Y198;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE Hypothetical protein OSUNBA0039018.17 (Hypothetical protein
DE OSUNBA0036M02.4).
GN Name=OSUNBA0039018.17; Synonyms=OSUNBA0036M02.4;
OS Oryza sativa (Japonica cultivar-group);
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzaceae; Oryza.
NCBI_TaxID=39947;
GN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tseltrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmiller S.B., Uteckback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (MAY-2003) to the EMBL/Genbank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Wang A., Maiti R., Jones K.M.,

RA Tallon L.J., Feldblum T.V., Tsitirin T., Bera J.J., Kim M.M., Jin S.,
 RA Fadrosh D., Vuong H., Overton II L.L., Reardon M., Weaver B.,
 RA Johri S., Uterback T.R., Pai G., Smith S., Wortman J., Haas B.J.,
 RA Zhu W., Yang Q., Koo H., Zismann V., Hsiao J., Quackenbush J.,
 RA White O., Salzberg S.L., Fraser C.M.,
 RT "Oryza sativa chromosome 3 BAC OSJNB0036M02 genomic sequence."
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell R.;
 RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC13930; AAP4639.1; -; Genomic DNA.
 DR EMBL; AC145388; AAP89142.1; -; Genomic DNA.
 DR Gremene; O7Y198; -;
 DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0016567; P:protein ubiquitination; IEA.
 DR InterPro; IPR001841; ZnF_ring.
 DR SMART; SMO0184; RING; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 319 AA; 33316 MW; FD51B063E2BAEB4 CRC64;
 Query Match 75.0%; Score 18; DB 2; Length 319;
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 OY 1 CXXXXXXC 8
 DB 296 CAAAAAAC 303
 RESULT 23
 ID 061A05_CABER PRELIMINARY; PRT; 320 AA.
 AC 061A05;
 DT 25-OCT-2004 (TReMBLrel. 28, Created)
 DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
 DE Hypothetical protein CBG13750.
 GN Name=CBG13750;
 OS Caenorhabditis briggsae.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6238;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RG The C.briggsae Sequencing Consortium;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL; CAAC01000066; CAB68107.1; -; Genomic DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004930; F:G-protein coupled receptor activity; IEA.
 DR InterPro; IPR000168; Nm7TM_chemrecept.
 KW Hypothetical protein.
 SQ SEQUENCE 320 AA; 36037 MW; 5E3389F6C7CF3C7B CRC64;
 Query Match 75.0%; Score 18; DB 2; Length 320;
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 OY 1 CXXXXXXC 8
 DB 302 CASSTSTC 309
 RESULT 24
 ID 062NM3_HUMAN PRELIMINARY; PRT; 328 AA.
 AC 062NM3;
 DT 05-JUL-2004 (TReMBLrel. 27, Created)

DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
 DE Hypothetical protein FLJ29006.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Stomach mucosa;
 RA Nimmiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
 RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
 RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Suzuki Y., Hata H.,
 RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,
 RA Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,
 RA Nagai K., Isogai T., Sugano S.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK131040; BAC85481.1; -; mRNA.
 SQ SEQUENCE 328 AA; 35381 MW; 7D04FC08E3270F8D CRC64;
 Query Match 75.0%; Score 18; DB 2; Length 328;
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 OY 1 CXXXXXXC 8
 DB 4 CTSASSSC 11

RESULT 25
 ID 062A86_ORYSA PRELIMINARY; PRT; 332 AA.
 AC 062A86;
 DT 05-JUL-2004 (TReMBLrel. 27, Created)
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
 DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
 DE Limonene cyclase like protein.
 GN Name=OSJNB0084107.20-2; Synonyms=OSJNB00002109.8-2;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsunoto T., Katayose Y.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 RT clone:OSJNB0084107.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsunoto T., Katayose Y.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 RT clone:OSJNB0002109.";
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005179; BAC93914.1; -; Genomic DNA.
 DR EMBL; AP005877; BAD31847.1; -; Genomic DNA.
 DR Gremene; Q624B6; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR InterPro; IPR004263; Exostosin.
 DR Pfam; PF03016; Exostosin; 1.
 SQ SEQUENCE 332 AA; 34451 MW; 57CDBBFC0CC5D6E CRC64;
 Query Match 75.0%; Score 18; DB 2; Length 332;
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
 DB 27 CAAAAAAC 34

```

RESULT 26
094589 9CILI PRELIMINARY; PRT: 350 AA.
ID 094589 9CILI PRELIMINARY; PRT: 350 AA.
AC 094589-
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Lemnabion factor.
OS Lemnabion bullinum.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peritritida;
OC Lemnabion.
OX NCBI_TaxID=54108;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Peters-Regehr T., Kusch J., Heckmann K.;
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Y09220; CA70420.1; -, Genomic_DNA.
DR HSSP; O16119; IEZG.
DR InterPro; IPR006212; Furin_repeat.
DR SMART; SM00261; FU; 3.
SQ SEQUENCE 350 AA; 35159 MW; DBR0C67654BD9D92E CRC64;

Query Match 75.0%; Score 18; DB 2; Length 350;
Beet Local Similarity 25.0%; Pred. No. 3.8e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
Db 105 CSTGATTC 112

RESULT 27
ID RHOM DROME STANDARD; PRT: 355 AA.
AC P20350; Q9W0F2;
DT 01-FEB-1991 (Rel. 17, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Rhomboid protein (EC 3.4.21.-) (Vainlet protein).
GN Name=rho; Synonyms=Ve; ORFNames=CG1004;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyridiidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=90249726; PubMed=2110920;
RA Bier E., Jan L.Y., Jan Y.N.;
RT "Rhomboid, a gene required for dorsoventral axis establishment and
peripheral nervous system development in Drosophila melanogaster.";
RL Genes Dev. 4:190-203(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STAIN=Beckley;
RA MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Stuton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abail J.F., Agbayant A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Berens P.V., Bernier B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
Butler J.K.C., Buesam D.A., Butler H., Cadieu L.B., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Dudin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,
Fowler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

```

```

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kerchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laško P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrlaks R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [3]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter U., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [4]
RP FUNCTION, AND MUTAGENESIS OF TRP-151; ARG-152; ASN-169; GLY-215;
RP SER-217 AND HIS-281.
RX MEDLINE=21526229; PubMed=11672525; DOI=10.1016/S0092-8674(01)00525-6;
RA Urban S., Lee J.R., Freeman M.;
RT "Drosophila Rhomboid-1 defines a family of putative intramembrane
serine proteases.";
RL Cell 107:173-182(2001).
CC -1- FUNCTION: Acts early in embryonic development to establish
position along the dorsoventral axis and then again later to
specify the fate of neuronal precursor cells. Involved in EGF
receptor signaling; cleaves Spitz to release the active growth
factor.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Golgi.
CC -1- DEVELOPMENTAL STAGE: Early blastoderm stages and later during
nervous development.
CC -1- SIMILARITY: Belongs to the peptidase S54 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC EMBL; X52454, CA36692.1; -, mRNA.
CC EMBL; AE003471, AA47496.1; -, Genomic_DNA.
CC PIR; A34597, A34597.
CC MEROPS; S54.001; -.
CC Ensembl; CG1004; Drosophila melanogaster.
CC FlyBase; FBgn0004635; rho.
CC GO; GO:0005794; C:Golgi apparatus; IDA.
CC GO; GO:0016021; C:Integral to membrane; TAS.
CC GO; GO:0008236; F:serine-type peptidase activity; IDA.
CC GO; GO:0007420; P:brain development; IMP.
CC GO; GO:0046845; P:branched duct epithelial cell fate determin. . .; TAS.
CC GO; GO:0001763; P:branching morphogenesis; NAS.
CC GO; GO:0007174; P:epidermal growth factor ligand processing; IGI.
CC GO; GO:0007479; P:leg disc proximal/distal pattern formation; TAS.
CC GO; GO:0007438; P:neocyte development; IGI.

```

DR GO; GO:0030707; P:ovarian follicle cell development (sensu In. . .; TAS.
 DR GO; GO:0007442; P:peripheral nervous system development; TAS.
 DR GO; GO:0045742; P:positive regulation of epidermal growth fac. . .; TAS.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; ICI
 DR GO; GO:0007166; P:regulation of epidermal growth factor recep. . .; TAS.
 DR GO; GO:0007432; P:salivary gland determination; NAS.
 DR GO; GO:0005202; P:tracheal sac formation (sensu Insecta); TAS.
 DR GO; GO:0007474; P:wing vein specification; NAS.
 DR InterPro: IPR002610; Rhomboid_like.
 DR Pfam: PF01694; Rhomboid_1.
 KM Developmental protein; Golgi stack; Hydrolyase; Protease;
 KW Serine protease; Transmembrane.
 FT TOPO_DOM 1 98 Cytoplasmic (Potential).
 FT TRANSMEM 99 119 Potential.
 FT TOPO_DOM 120 162 Luminal (Potential).
 FT TRANSMEM 163 183 Potential.
 FT TOPO_DOM 184 188 Cytoplasmic (Potential).
 FT TRANSMEM 189 209 Potential.
 FT TOPO_DOM 210 210 Luminal (Potential).
 FT TRANSMEM 211 231 Potential.
 FT TOPO_DOM 232 244 Cytoplasmic (Potential).
 FT TRANSMEM 245 265 Potential.
 FT TOPO_DOM 266 275 Luminal (Potential).
 FT TRANSMEM 276 296 Potential.
 FT TOPO_DOM 297 308 Cytoplasmic (Potential).
 FT TRANSMEM 309 329 Potential.
 FT TOPO_DOM 330 355 Luminal (Potential).
 FT ACT_SITE 169 169 Charge relay system.
 FT ACT_SITE 217 217 Charge relay system.
 FT ACT_SITE 217 217 Charge relay system.
 FT ACT_SITE 281 281 Charge relay system.
 FT MUTAGEN 151 151 W->A: Abolishes protease activity.
 FT MUTAGEN 152 152 R->A: Abolishes protease activity.
 FT MUTAGEN 169 169 N->A: Abolishes protease activity.
 FT MUTAGEN 215 215 G->A: Abolishes protease activity.
 FT MUTAGEN 217 217 S->C,T: Abolishes protease activity.
 FT MUTAGEN 281 281 H->A: Abolishes protease activity.
 FT CONFLICT 4 4 L -> P (in Ref. 1).
 FT CONFLICT 33 33 A -> V (in Ref. 1).
 FT CONFLICT 46 46 S -> T (in Ref. 1).
 SQ SEQUENCE 355 AA; 39330 MW; 2E0572D08831C5D0 CRC64;

Query Match 75.0%; Score 18; DB 1; Length 355;
 Best Local Similarity 25.0%; Pred. No. 3.8e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
 DB 57 CSTASSTC 64

RESULT 28
 OS40V7 DROME PRELIMINARY; PRT; 355 AA.
 ID OS40V7;
 AC OS40V7;
 DT 13-SEP-2005 (Tremblrel. 31, Created)
 DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
 DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
 DE LDD6131P.
 GN Name=rho;
 OS Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OK NCBI_Taxid=7227;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Berkeley.
 RA Stapleton M., Brockslein P., Hong L., Aghayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresner D., Farfan D., Friese B.,
 RA George R., Gonzalez W., Guarin H., Krommiller B., Li P., Liso G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Patagas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Ceiniker S.;

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY121641; AAM51968.1; -, mRNA.
 SQ SEQUENCE 355 AA; 39330 MW; 2E0572D08831C5D0 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 355;
 Best Local Similarity 25.0%; Pred. No. 3.8e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
 DB 57 CSTASSTC 64

RESULT 29
 TISD_MOUSE STANDARD; PRT; 367 AA.
 ID TISD_MOUSE
 AC P23949;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Butyrate response factor 2 (TIS1D protein).
 GN Name=zfp3612; Synonyms=Brf2, Tis1d;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OC NCBI_Taxid=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=BALB/c;
 RX MEDLINE=91141531; PubMed=1996120;
 RA Varum B.C., Ma Q., Chi T., Fletcher B., Herschman H.R.;
 RT "The TIS1 primary response gene is a member of a gene family that
 RT encodes proteins with a highly conserved sequence containing an
 RT unusual Cys-His repeat.";
 RL Mol. Cell. Biol. 11:1754-1758(1991).
 RN [2]
 RP NUCLEOTIDE SEQUENCE OF 1-50.
 RA Fletcher B.S.;
 RL Thesis (1992); University of California Los Angeles, United States.
 CC -1- FUNCTION: Probable regulatory protein involved in regulating the
 CC response to growth factors.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- SIMILARITY: Contains 2 C3H1-type zinc fingers.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

DR EMBL; M58564; AAA72946.1; -, mRNA.
 DR EMBL; M97165; AAA39709.1; -, Genomic_DNA.
 DR PIR; C39590; C39590.
 DR HSSP; P22893; 1M90.
 DR SMK; P23949; 124-193.
 DR Ensemble; ENSMUSG00000045817; Mus musculus.
 DR MGI; MGI:107945; Zfp3612.
 DR InterPro; IPR007635; TIS1B_N.
 DR InterPro; IPR000571; Znf_CCH.
 DR Pfam; PF04553; Tis1B_C_1.
 DR Pfam; PF00642; zf-CCH1_2.
 DR SMART; SM00356; Znf_CCH1_2.
 DR DNA-binding; Metal-Binding; Nuclear protein; Repeat; Zinc;
 KW Zinc-finger.
 FT ZN_FING 132 151 C3H1-type 1.
 FT ZN_FING 170 189 C3H1-type 2.
 FT COMPBIAS 64 67 Poly-Gly.
 FT COMPBIAS 111 114 Poly-Gly.
 FT COMPBIAS 200 203 Poly-Gly.
 FT COMPBIAS 263 266 Poly-Pro.
 FT COMPBIAS 297 303 Poly-Ala.
 SQ SEQUENCE 367 AA; 37593 MW; 361244AF6244E46E CRC64;

Query Match 75.0%; Score 18; DB 1; Length 367;
 Best Local Similarity 25.0%; Pred. No. 3.9e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
 Db 273 CSSSASGC 280

RESULT 30

086SB6_DROVI PRT; 404 AA.
 AC 086SB6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Rhomboid.
 GN Name=rho;
 OS Drosophila virilis (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 NX NCBI_TaxID=7244;
 RX MEDLINE=22844358; PubMed=12963110; DOI=10.1016/S0925-4773(03)00164-3;
 RN NCLEBOTIDE SEQUENCE.

RA Nakamura Y., Matsuno K.;
 RT "Species-specific activation of EGF receptor signaling underlies
 RT evolutionary diversity in the dorsal appendage number of the genus
 RT Drosophila egsheila.";
 RL Mech. Dev. 120:897-907(2003).
 DR EMBL; AB089248; BAC56701.1; -; mRNA.
 DR MEROPS; S54.001; -;
 DR FlyBase; FBgn0062278; Dv|Flyb|rho.
 DR InterPro; IPR002610; Rhomboid-like.
 DR Pfam; PF01694; Rhomboid; 1.
 SQ SEQUENCE 404 AA; 44035 MW; EF641632B02F9011 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 404;
 Best Local Similarity 25.0%; Pred. No. 4e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
 Db 63 CSTRASSTC 70

RESULT 31
 06ZAE5_ORYSA PRELIMINARY; PRT; 417 AA.
 AC 06ZAE5;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Limonene cyclase like protein.
 GN Name=OSJNB0084L07.20-3; Synonyms=OSJNB0002L09.8-3;
 OS Oryza sativa (japonica cultivar-group).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Erihartoideae; Oryzae; Oryza.
 NX NCBI_TaxID=39947;
 RX MEDLINE=39947;
 RN NCLEBOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbata (G43) genomic DNA, chromosome 7, BAC
 RT clone:OSJNB0084L07.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NCLEBOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbata (G43) genomic DNA, chromosome 7, BAC
 RT clone:OSJNB002L09.";

RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP005179; BAC3915.1; -; Genomic DNA.
 DR EMBL; AP005877; BAC31849.1; -; Genomic DNA.
 DR Gramene; 06ZAE5; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR InterPro; IPR004263; Exostosin.
 DR Pfam; PF03016; Exostosin; 1.
 SQ SEQUENCE 417 AA; 44610 MW; 676689A53BIDC6A CRC64;

Query Match 75.0%; Score 18; DB 2; Length 417;
 Best Local Similarity 25.0%; Pred. No. 4.1e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
 Db 27 CAANAAC 34

RESULT 32

08UTL4_ORYLA PRELIMINARY; PRT; 445 AA.
 ID 08UTL4_ORYLA
 AC 08UTL4;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE TAPBP Protein.
 GN Name=TAPBP;
 OS Oryzias latipes (Japanese ricefish).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 CC Belontiiformes; Adrianchthyidae; Oryziinae; Oryzias.
 NX NCBI_TaxID=8090;
 RX MEDLINE=21850510; PubMed=11862394; DOI=10.1007/s00251-001-0427-3;
 RN NCLEBOTIDE SEQUENCE.

RA Matsuo M.Y., Asakawa S., Shimizu N., Kimura H., Nonaka M.;
 RT "Nucleotide sequence of the MHC class I genomic region of a teleost,
 RT the medaka (Oryzias latipes).";
 RL Immunogenetics 53:930-940(2002).
 DR EMBL; BA000027; BAB83851.1; -; Genomic DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0019885; P:antigen processing, endogenous antigen via . . .; IEA.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR008056; Tapasin.
 DR Pfam; PF00047; IG; 1.
 DR PRINTS; PRO1669; TAPASIN.
 DR SMART; SM00409; IG; 2.
 DR PROSITE; PSS0835; IG LIKE; 2.
 KW Immunoglobulin domain; Repeat.
 SQ SEQUENCE 445 AA; 48260 MW; FDF7C5FBE74370BB CRC64;

Query Match 75.0%; Score 18; DB 2; Length 445;
 Best Local Similarity 25.0%; Pred. No. 4.2e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
 Db 22 CSSSSSSC 29

RESULT 33

07SHC6_NEUCR PRELIMINARY; PRT; 480 AA.
 ID 07SHC6_NEUCR
 AC 07SHC6;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Predicted protein (Hypothetical protein B13N4.210).
 GN Name=NCU01880.1; Synonyms=B13N4.210;
 OS Neurospora crassa.

CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 CC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 CC NCBI_TaxID=5141;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=OR74A;
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker J.E., Read N.D.,
 RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Reiman B.,
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
 RA Qui D., Janakiev P., Pedersen D., Nelson M., Washburne M.,
 RA Seltreknoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
 RA Koche G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
 RA Kamal M., Kamysella M., Mauceli E., Bielke C., Rudd S., Frisken D.,
 RA Kyrstova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken D.,
 RA Cogoni C., Marino G., Catchside D., Li W., Pratt R.U., Osmann S.A.,
 RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
 RA Yarden O., Plamann M., Siller S., Dunlap J., Radford A., Aramayo R.,
 RA Nativig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
 RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
 RL Nature 0:0-0(2003).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Schulte U., Aign V., Hohenseil J., Brandt P., Fartmann B., Holland R.,
 RA Nykatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (DEC-2003) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RA German Neurospora genome project;
 RL Submitted (DEC-2003) to the EMBL/Genbank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL; AABX01000004; EAA36275.1; -; Genomic_DNA.
 CC EMBL; BX842681; CAE81980.1; -; Genomic_DNA.
 CC Hypothetical protein.
 SQ SEQUENCE 480 AA; 50449 MW; E073D8EBC33C9A2B CRC64;

Query Match 75.0%; Score 18; DB 2; Length 480;
 Best Local Similarity 25.0%; Pred. No. 4.3e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
 Db 127 CSSSSSSC 134

RESULT 34
 0758Y3 ASHGO PRELIMINARY; PRT; 480 AA.
 AC 0758Y3;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE AD3395CD.
 GN Name=ADR395C;
 OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetidae; Saccharomycetaceae; Eremothecium.
 CC NCBI_TaxID=33169;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=ATCC 10895;
 RX PubMed=15001715; DOI=10.1126/science.1095781;
 RA Dietrich F.S., Voegel S., Brachat S., Lerch A., Gates K., Steiner S.,
 RA Morf C., Boehmann R., Iuedi P., Choi S., Wing R.A., Flavier A.,
 RA Gaffney T.D., Philippson P.;
 RT "The Ashbya gossypii genome as a tool for mapping the ancient
 RT Saccharomycetes cerevisiae genome."
 RL Science 304:304-307(2004).
 DR EMBL; AE016817; AAS52314.1; -; Genomic_DNA.
 DR ACD; ADR395C; -.

DR GO; GO:000414; F:3', 5'-cyclic-nucleotide phosphodiesterase a. .; IEA.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR003607; Met_phos_hydro.
 DR InterPro; IPR02073; PDBase.
 DR Pfam; PF00233; PDBase_1; 1.
 DR PRINTS; PR00387; PDIESTERASE1.
 DR SMART; SM00471; Hdc; 1.
 DR PROSITE; PS00126; PDBASE_1; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 480 AA; 53835 MW; 75AB389455D01CB CRC64;

Query Match 75.0%; Score 18; DB 2; Length 480;
 Best Local Similarity 25.0%; Pred. No. 4.3e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
 Db 205 CTAATAC 212

RESULT 35
 T1SD HUMAN STANDARD; PRT; 492 AA.
 AC P47974; Q9BSJ3;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Butyrate response factor 2 (T1SID protein) (BGF-response factor 2)
 DE (ERF-2).
 GN Name=ZFP3612; Synonyms=BRF2, ERF2, T1SID;
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 CC Homo.
 CC NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=96132724; PubMed=8545129;
 RA Ino T., Yasui H., Hirano M., Kurosawa Y.;
 RT "Identification of a member of the T1SID early response gene family at
 RT the insertion point of a DNA fragment containing a gene for the T-cell
 RT receptor beta chain in an acute T-cell leukemia."
 RL Oncogene 11:2705-2710(1995).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=95137407; PubMed=7835719; DOI=10.1016/0378-1119(94)00636-P;
 RA Nie X.F., Maclean K.N., Kumar V., McKay I.A., Bustin S.A.;
 RT "ERF-2, the human homologue of the murine T1sid early response
 RT gene."
 RL Gene 152:285-286(1995).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Pancreas;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uesdi T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Lochellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green B.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Maira M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Probable regulatory protein involved in regulating the
CC response to growth factors.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: Contains 2 C3H1-type zinc fingers.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; U07802; AAA91778.1; -; Genomic_DNA.
DR EMBL; X78992; CAAS5592.1; -; mRNA.
DR EMBL; BC005010; AAH05010.1; -; mRNA.
DR PIR; S49147; S49147.
DR PDB; 1RGO; NMR; A:151-220.
DR Ensembl; ENSG00000152518; Homo sapiens.
DR HGNC; HGNC:1108; ZFP36L2.
DR H-InvDB; HIX0002007; -.
DR GO; GO:0003700; F:transcription factor activity; TAS.
DR GO; GO:0008283; P:cell proliferation; TAS.
DR InterPro; IPR007635; Tis11B_N.
DR InterPro; IPR005571; Znf_CCH.
DR Pfam; PF04553; Tis11B_C; 1.
DR Pfam; PF00642; zf-CCCH; 2.
DR SMART; SM00356; ZNF_C3H1; 2.
KW 3D-structure; DNA-binding; Metal-binding; Nuclear protein; Repeat;
KW Zinc; Zinc-finger.
FT ZN_FING 159 178 C3H1-type 1.
FT ZN_FING 197 216 C3H1-type 2.
FT COMPBIAS 106 109 Poly-Gly.
FT COMPBIAS 138 141 Poly-Gln.
FT COMPBIAS 143 146 Poly-Gly.
FT COMPBIAS 288 291 Poly-Pio.
FT COMPBIAS 323 330 Poly-Ala.
FT COMPBIAS 382 388 Poly-Gln.
FT COMPBIAS 393 399 TS -> DL (in Ref. 1).
FT CONFLICT 96 97 AA -> T (in Ref. 2).
FT CONFLICT 318 318 AA -> LR (in Ref. 2).
FT CONFLICT 329 330 A -> AAA (in Ref. 3).
FT CONFLICT 330 330 Q -> QQQQ (in Ref. 3).
FT CONFLICT 399 399 Missing (in Ref. 1).
SQ SEQUENCE 492 AA; 50921 MW; E4E2EE26691CCF7 CRC64;
Query Match 75.0%; Score 18; DB 1; Length 492;
Best Local Similarity 25.0%; Pred. No. 4.4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CXXXXXXC 8
DB 299 CSSASASC 306

RESULT 36
053TB4 HUMAN
ID 053TB4_HUMAN PRELIMINARY; PRT; 494 AA.
AC 053TB4;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Hypothetical protein ZFP36L2.
GN Name=ZFP36L2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Abbott A., Boyer E., Heyen J.;

RT "The sequence of Homo sapiens BAC clone RP11-339H12.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.H.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC010883; AA14992.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 494 AA; 51062 MW; 10E23FA9E2DDABD4 CRC64;
Query Match 75.0%; Score 18; DB 2; Length 494;
Best Local Similarity 25.0%; Pred. No. 4.4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CXXXXXXC 8
DB 299 CSSASASC 306

RESULT 37
06PD21 MOUSE
ID 06PD21_MOUSE PRELIMINARY; PRT; 526 AA.
AC 06PD21;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Shd protein (Fragment).
GN Name=Shd;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marisina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Tohiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.U., McKernan K.J., Malek J.A., Gunnaracne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Scherach A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strauberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC058986; AAH58986.1; -; mRNA.
DR HSSP; P00524; IKC2.
DR GO; GO:0005515; F:protein binding; IDA.

DR GO:0005070; F:SH3/SH2 adaptor activity; TAS.
 DR GO:0042100; P:B-cell proliferation; IDA.
 DR InterPro: IPR00980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRODOM: PD00093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS0001; SH2; 1.
 FT NON TER
 SQ SEQUENCE 526 AA; 57393 MW; 6F358E391E2C63E0 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 526;
 Best Local Similarity 25.0%; Pred. No. 4.5e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
 Db 71 CSASSASC 78

RESULT 38
 OSB195_DROME PRELIMINARY; PRT; 544 AA.
 ID Q5B195;
 AC Q5B195;
 DT 10-MAY-2005 (TREMBLrel. 30, Created)
 DT 10-MAY-2005 (TREMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
 DE RE04051P.
 GN Name=dpa;
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydriidae; Drosophilidae; Drosophila.
 CC NCBI_TaxID=7227;
 OK [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=Berkley;
 RA Stapleton M., Carlson J., Chavez C., Friese B., George R., Pacleb J.,
 RA Pak S., Wan K., Yu C., Rubin G.M., Celitker S.,
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BT021329; AAX3477.1; -; mRNA.
 DR GO:0005634; C:nucleus; IEA.
 DR GO:0005524; F:ATP binding; IEA.
 DR GO:0003677; F:DNA binding; IEA.
 DR GO:0008094; F:DNA replication; IEA.
 DR GO:0006270; P:DNA replication initiation; IEA.
 DR InterPro: IPR01304; Lectin_C.
 DR InterPro: IPR01208; MCM_4.
 DR InterPro: IPR008047; MCM_4.
 DR PRINTS: PR01660; MCMPROTEIN4.
 DR SMART: SM00350; MCM; 1.
 DR PROSITE: PS00615; C-TYPE LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS50051; MCM_2; 1.
 SQ SEQUENCE 544 AA; 60347 MW; 8D221F998960D707 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 544;
 Best Local Similarity 25.0%; Pred. No. 4.5e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
 Db 522 CSTSTSC 529

RESULT 39
 OS1TR2_TUHV1
 ID Q91TR2; TUHV1 PRELIMINARY; PRT; 545 AA.
 AC Q91TR2;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE T31.
 SQ Tupaiid herpesvirus 1 (strain 1) (TuHV-1) (Herpesvirus tupai) (strain

OS 1)).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae.
 CC NCBI_TaxID=10397;
 OK [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=2;
 RX MEDLINE=21211637; PubMed=11312357;
 RX DOI=10.1128/JVI.75.10.4854-4870.2001;
 RA Bahr U., Darai G.;
 RT "Analysis and characterization of the complete genome of tupai (tree
 RT shrew) herpesvirus."
 RT J. Virol. 75:4854-4870(2001).
 DR EMBL; AF281817; AAK57075.1; -; Genomic_DNA.
 DR InterPro: IPR007578; DUF570.
 DR Pfam: PF04489; DUF570; 1.
 SQ SEQUENCE 545 AA; 59508 MW; 98B5EC0DEF7A806 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 545;
 Best Local Similarity 25.0%; Pred. No. 4.5e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
 Db 535 CAASATC 542

RESULT 40
 OS28N9_ORYSA
 ID Q5Z8N9;
 AC Q5Z8N9;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE Aspartic proteinase nepenthesin II-like.
 GN Name=P0541C02.19-2;
 OS Oryza sativa (japonica cultivar-group).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Ehrhacoidae; Oryzae; Oryza.
 CC NCBI_TaxID=39947;
 OK [1]
 RN NUCLEOTIDE SEQUENCE.
 RP Sasaki T., Matsumoto T., Yamamoto K.;
 RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
 RT clone: P0541C02.19-2" to the EMBL/GenBank/DBJ databases.
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003769; BAD61723.1; -; Genomic_DNA.
 DR GO:0004194; F:pepsin A activity; IEA.
 DR GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR01461; Peptidase A1.
 DR InterPro: IPR01969; Pept_Asp_AS.
 DR Pfam: PF00026; Asp; 1.
 DR PRINTS: PR00792; PEPsin.
 DR PROSITE: PS00141; ASP PROTEASE; UNKNOWN_2.
 SQ SEQUENCE 551 AA; 57586 MW; BE78840EB14824DF CRC64;

Query Match 75.0%; Score 18; DB 2; Length 551;
 Best Local Similarity 25.0%; Pred. No. 4.6e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
 Db 182 CATATSSC 189

RESULT 41
 OSB30_ORYSA
 ID Q8SB30; ORYSA PRELIMINARY; PRT; 551 AA.
 AC Q8SB30;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

DE Putative chloroplast nucleoid DNA-binding protein.
 GN Name=OJ1540.H01.13;
 OS Oryza sativa (Rice).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Euphorbiaceae; Oryzae; Oryza.
 CC NCBI_TaxID=4530;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell C.R., Yuan O., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
 RA Gansberger K., Bremner M., Burgess S., Hance M., Shvartsbeyn M.,
 RA Tstirna T., Riggs F., Hsiao J., Zismann V., Blunt S., Pal G.,
 RA Valachen S.E., Ultebrack T.R., Feldblum T.V., Kalb E., Quackenbush J.,
 RA Salzberg S.L., White O., Fraser C.M.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC091774; AAL79734.1; -; Genomic_DNA.
 DR HSSP; P00797; 2REN.
 DR Gramene; O8S310; -;
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0004194; F:pepsin A activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001461; Peptidase A1.
 DR InterPro; IPR001969; Pept_Asp_AS.
 DR Pfam; PF00026; Asp_1.
 DR PRINTS; PR00792; PEPsin.
 DR PROSITE; PS00141; ASP_PROTASE; UNKNOWN_2.
 DR DNAS-binding.
 KM SEQUENCE 551 AA; 57586 MW; BE78840EB14824DF CRC64;
 SQ
 Query Match 75.0%; Score 18; DB 2; Length 551;
 Best Local Similarity 25.0%; Pred. No. 4.6e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CXXXXXXC 8
 DB 182 CATATSSC 189

RESULT 42
 Q5AP57 CANAL
 ID Q5AP57 CANAL PRELIMINARY; PRT; 552 AA.
 AC Q5AP57;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Hypothetical yfw family protein 5.
 GN ORFNames=CaO19.4881;
 OS Candida albicans SC5314.
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; mitospotic Saccharomycetales; Candida.
 CC NCBI_TaxID=237561;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA STRAIN=SC5314;
 RX PubMed=15123810; DOI=10.1073/pnas.0401648101;
 RA Jones T., Federpiel N.A., Chibana H., Dungan J., Kalman S.,
 RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
 RA Davis R.W., Scherer S.;
 RT "The diploid genome sequence of Candida albicans";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334 (2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA STRAIN=SC5314;
 RA Dungan J., Kuo A., Newport G., Lan C.-Y., Iijima C., Adegoola O.,
 RA Roberts J., Petersen K., Donnelly S., Favoreto S., Tzung K.-W.,
 RA Jones T., Scherer S., Agabian N.;
 RT "Annotation of the Genome of Candida albicans";
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAC01000001; EAL04869.1; -; Genomic_DNA.
 KM Hypothetical protein.
 SQ SEQUENCE 552 AA; 62256 MW; 5CA72B0ED5B87B3F CRC64;
 Query Match 75.0%; Score 18; DB 2; Length 552;
 Best Local Similarity 25.0%; Pred. No. 4.6e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CXXXXXXC 8
 DB 466 CTSATTSC 473

RESULT 43
 Q80J53 HPBV0
 ID Q80J53 HPBV0 PRELIMINARY; PRT; 560 AA.
 AC Q80J53;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Truncated polymerase.
 OS Hepatitis B virus.
 CC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 CC NCBI_TaxID=10407;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Lin X., Xu X., Zheng D.;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY238972; AAC64453.1; -; Genomic_DNA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR001462; P:RNA-dependent DNA replication; IEA.
 DR InterPro; IPR000201; DNAPol_viral_C.
 DR InterPro; IPR00477; RTase.
 DR Pfam; PF00336; DNA_pol_viral_C; 1.
 DR Pfam; PF00242; DNA_pol_viral_N; 1.
 DR Pfam; PF00078; RVT_1; 1.
 SQ SEQUENCE 560 AA; 62448 MW; 7BCD0DF6F891CE4B CRC64;
 Query Match 75.0%; Score 18; DB 2; Length 560;
 Best Local Similarity 25.0%; Pred. No. 4.6e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CXXXXXXC 8
 DB 268 CASSSSC 275

RESULT 44
 Q6N1S1 CORDI
 ID Q6N1S1 CORDI PRELIMINARY; PRT; 581 AA.
 AC Q6N1S1;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Putative secreted protein.
 GN OrderedlocusNames=DIP0696;
 OS Corynebacterium diphtheriae.
 CC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 CC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 CC NCBI_TaxID=1717;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA STRAIN=Biotype gravis / NCTC 13129;
 RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/9kg874;
 RA Cerdano-Tarraga A.-M., Efstratiou A., Dover L.G., Holden M.T.G.,

RA Pallen M.J., Bentley S.D., Beara G.S., Churcher C.M., James K.D.,
 RA de Zorja A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
 RA Hamlin N., Holtroyd S., Jagels K., Moule S., Quail M.A.,
 RA Rabinowitch E., Rutherford K.M., Thomson N.R., Unwin L.,
 RA Whitehead S., Barrett B.G., Parkhill J.,
 RT "The complete genome sequence and analysis of *Corynebacterium*
 RT *diphtheriae* NCTC13129."
 RL Nucleic Acids Res. 31:6516-6523 (2003).
 DR EMBL: BX248355; CAB49213.1; -; Genomic_DNA.
 KW Complete proteome.
 SQ SEQUENCE 581 AA; 62851 MW; 9CF98BC85A311AC3 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 581;
 Best Local Similarity 25.0%; Pred. No. 4.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
 DB 17 CAATTAAC 24

RESULT 45
 Q51VZ0_MAGGR PRELIMINARY; PRT; 592 AA.
 AC Q51VZ0;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=MG03395.4;
 OS Magnaporthe grisea 70-15.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaportha.
 OX NCBI_TaxID=242507;
 RX [1]
 NUCLEOTIDE SEQUENCE.
 RA Birren B., Nisbaum C., Abebe A., Abouelleil A., Adekoya E.,
 RA Alt-zehra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
 RA Archchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
 RA Bayul T., Bilshteyn B., Bloom T., Blye J., Boguslavskiy L.,
 RA Bowdery M., Bouhgalter B., Brunache A., Butler J., Calixte N.,
 RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
 RA Collimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
 RA David R., Dawe T., Degray S., Dodge S., Dooley K., Dorje P.,
 RA Dorjee K., Dorris L., Dufley N., Dupes A., Ekins T., Engels R.,
 RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
 RA Fitzgerald M., Foley K., Gage D., Galagan J., Geatin G., Gnerre S.,
 RA Gutirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
 RA Homan T., Horn A., Houde N., Hughes L., Hulme W., Hubby E., Iliev I.,
 RA Jaffe D., Jones C., Kamil M., Kamat A., Kamysseils M., Karlsson E.,
 RA Kells C., Kien A., Kisher P., Kodira C., Kulbokas E., Labutti K.,
 RA Lama D., Landers T., Leger P., Levine S., Lewis D., Lewis T.,
 RA Lindblad-Toh K., Liu X., Loksitsang T., Loksitsang Y., Lucien O.,
 RA Lui A., Ma L.J., Mahbit R., Macdonald J., Maclean C., Major J.,
 RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
 RA McCarthy M., McDonough S., McGhee T., Meldrum J., Menus L.,
 RA Meisner J., Mihalev A., Minova T., Mikkelson T., Mienna V., Moru K.,
 RA Moses J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
 RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizari M., Norby C.,
 RA Norbu N., O'Donnell P., Okawa O., O'Leary S., Omotohio B.,
 RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
 RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
 RA Reta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
 RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
 RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnuez C.,
 RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
 RA Stetson K., Stone C., Stone S., Stubbs M., Talmas J., Thunga P.,
 RA Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
 RA Towey S., Tsamla T., Tesomo N., Vallee D., Vassiliev H.,
 RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
 RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,

RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
 RA Zimmer A., Zody M., Lander E.,
 RT "The genome sequence of *Magnaporthe grisea*."
 RT Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Dean R., Mitchell T., Brown D., Pan H., Thon M.,
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.,
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 CC -! CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL: AACU01001048; EAA50236.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 592 AA; 63027 MW; 84D30AEC5A5D8ABC CRC64;

Query Match 75.0%; Score 18; DB 2; Length 592;
 Best Local Similarity 25.0%; Pred. No. 4.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
 DB 347 CAAASTGC 354

RESULT 46
 Q6Z4E7_ORYSA PRELIMINARY; PRT; 606 AA.
 AC Q6Z4E7;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Putative pectin-glucuronyltransferase.
 GN Name=OSUNB0084107.20-1; Synonyms=OSUNB0002109.8-1;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RX [1]
 NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.,
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 RT clone:OSUNB0084107."
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.,
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 RT clone:OSUNB0002109."
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP005179; BAC3913.1; -; Genomic_DNA.
 DR EMBL: AP005877; BAC1848.1; -; Genomic_DNA.
 DR Gramene; O6Z4E7; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR004263; Exoscosin.
 DR Pfam; PF03016; Exoscosin; 1.
 KW Transferase.
 SQ SEQUENCE 606 AA; 65830 MW; F90248B809A18266 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 606;
 Best Local Similarity 25.0%; Pred. No. 4.7e+02;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8
 DB 27 CAAAAAAC 34

```

RESULT 47
QARNV9_TETNG PRELIMINARY; PRT; 691 AA.
ID QARNV9_TETNG PRELIMINARY; PRT; 691 AA.
AC QARNV9_TETNG PRELIMINARY; PRT; 691 AA.
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Chromosome 10 SCAFI5009, whole genome shotgun sequence.
GN ORFNames=ESTENG00031390001;
OS Tetradon nigraviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neocelostei;
OC Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OC NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dosset C., Segurens B.,
RA Daelva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthonard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Catcolico L., Poullain J., De Bernardis V.,
RA Cruaud C., Duprat S., Brotier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chappe C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Landet V., Schachter V., Queller F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetraodon nigraviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAPE01015009; CAG09923.1; -; Genomic DNA.
SQ SEQUENCE 691 AA; 78163 MW; 440CAE5E405255CD CRC64;

Query Match 75.0%; Score 18; DB 2; Length 691;
Best Local Similarity 25.0%; Pred. No. 5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
DB 288 CXXXXXXC 295

RESULT 48
Q7S136_NEUCR PRELIMINARY; PRT; 705 AA.
ID Q7S136_NEUCR PRELIMINARY; PRT; 705 AA.
AC Q7S136_NEUCR PRELIMINARY; PRT; 705 AA.
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU09990.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OC NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., Fitzhugh W., Ma L.-U., Smirnov S., Purcell S., Reiman B.,
RA Elting T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,

```

```

RA Selltremlkoff C.P., Kinsey J.A., Braut E.L., Zelter A., Schulte U.,
RA Koche G.O., Jedd G., Mewes W., Straben G., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysseilis M., Mauceli E., Bielke C., Ruda S., Frisman D.,
RA Kryzstofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmari S.A.,
RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Piamann M., Sella S., Dunlap J., Radford A., Aramayo R.,
RA Narvig D.O., Alex L.A., Mannhaupt G., Rhohe D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,
RT "The genome sequence of the filamentous fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX0100465; EAA29063.1; -; Genomic DNA.
SQ SEQUENCE 705 AA; 75444 MW; D74E0BF0243E134 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 705;
Best Local Similarity 25.0%; Pred. No. 5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
DB 367 CXXXXXXC 374

RESULT 49
Q5RFW2_BRABE PRELIMINARY; PRT; 711 AA.
ID Q5RFW2_BRABE PRELIMINARY; PRT; 711 AA.
AC Q5RFW2_BRABE PRELIMINARY; PRT; 711 AA.
DT 01-FEB-2005 (TREMBlrel. 29, Created)
DT 01-FEB-2005 (TREMBlrel. 29, Last sequence update)
DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
DE Novel protein (Zgc:77446).
GN Name=CH211-200P13.3; ORFNames=CH211-200P13.3-001, zgc:77446;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Donaldson S.;
RL Submitted (DEC-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL; CR536610; CAIL1747.1; -; Genomic DNA.
DR ZFIN; ZDB-GENE-040426-933; zgc:77446.
DR InterPro; IPR001876; Znf_RandP.
DR InterPro; IPR001876; Znf_RandP.
DR Pfam; PF02845; CUE; 1.
DR SMART; SM00546; CUE; 1.
DR SMART; SM00547; ZNF_RAZ; 1.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS50199; ZF_RANBP2_2; 1.
SQ SEQUENCE 711 AA; 76880 MW; E26723DF30859356 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 711;
Best Local Similarity 25.0%; Pred. No. 5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
DB 330 CXXXXXXC 337

RESULT 50
Q5AP84_CANAL PRELIMINARY; PRT; 722 AA.
ID Q5AP84_CANAL PRELIMINARY; PRT; 722 AA.
AC Q5AP84_CANAL PRELIMINARY; PRT; 722 AA.
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
DE Hypothetical yfw family protein 5.

```

GN Name=YFW5; ORFNames=CaO19.12344;
OS Candida albicans SC5314.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=237561;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
RX PubMed=15123810; DOI=10.1073/pnas.0401648101;
RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,
RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
RA Davis R.W., Scherer S.;
RT "The diploid genome sequence of Candida albicans."
RL Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
RA Dungan J., Kuo A., Newport G., Ian C.-Y., Iijima C., Adegbole O.,
RA Roberts J., Persson K., Donnelly S., Favoreto S., Tzung K.-W.,
RA Jones T., Scherer S., Agabian N.;
RT "Annotation of the Genome of Candida albicans."
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAC001000002; EAL04673.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 722 AA; 81770 MW; 16E7C88A35A416AE CRC64;

Query Match 75.0%; Score 18; DB 2; Length 722;
Best Local Similarity 25.0%; Pred. No. 5.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXC 8
Db 636 CTSARTSC 643

Search completed: January 4, 2006, 16:09:52
Job time : 54.2435 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 ; Search time 61.5913 Seconds
(without alignments)
64.204 Million cell updates/sec

Title: US-09-932-322-10

Perfect score: 25

Sequence: 1 CXXXXXXC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :

1: A_Geneseq_21.*
2: geneseqp19808:*
3: geneseqp20005:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20038:*
8: geneseqp20048:*
9: geneseqp20058:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	72.0	59	4	ABBA1147 Peptide #
2	18	72.0	59	4	ABBA1147 Peptide #
3	18	72.0	59	4	ABBA1147 Peptide #
4	18	72.0	59	4	ABBA1147 Peptide #
5	18	72.0	59	4	ABBA1147 Peptide #
6	18	72.0	59	4	ABBA1147 Peptide #
7	18	72.0	59	4	ABBA1147 Peptide #
8	18	72.0	59	4	ABBA1147 Peptide #
9	18	72.0	59	4	ABBA1147 Peptide #
10	18	72.0	59	4	ABBA1147 Peptide #
11	18	72.0	59	4	ABBA1147 Peptide #
12	18	72.0	59	4	ABBA1147 Peptide #
13	18	72.0	59	4	ABBA1147 Peptide #
14	18	72.0	59	4	ABBA1147 Peptide #
15	18	72.0	59	4	ABBA1147 Peptide #
16	18	72.0	59	4	ABBA1147 Peptide #
17	18	72.0	59	4	ABBA1147 Peptide #
18	18	72.0	59	4	ABBA1147 Peptide #
19	18	72.0	59	4	ABBA1147 Peptide #
20	18	72.0	59	4	ABBA1147 Peptide #
21	18	72.0	59	4	ABBA1147 Peptide #
22	18	72.0	59	4	ABBA1147 Peptide #
23	18	72.0	59	4	ABBA1147 Peptide #
24	18	72.0	59	4	ABBA1147 Peptide #

25	18	72.0	169	7	ABO83337 Pseudomon
26	18	72.0	174	7	ADP30801 Human sec
27	18	72.0	193	7	ABO73977 Pseudomon
28	18	72.0	198	8	ADP30493 Human sec
29	18	72.0	198	8	ADP30492 Human sec
30	18	72.0	198	8	ADP30477 Human sec
31	18	72.0	198	8	ADP30481 Human sec
32	18	72.0	198	8	ADP30491 Human sec
33	18	72.0	228	8	ADP31281 Human sec
34	18	72.0	229	7	ABO80501 Pseudomon
35	18	72.0	233	7	ABO77555 Pseudomon
36	18	72.0	234	8	ADP31468 Human sec
37	18	72.0	242	8	ADP31616 Human sec
38	18	72.0	246	5	AAU75930 Androgen
39	18	72.0	246	7	ABO74848 Pseudomon
40	18	72.0	249	8	ADP30754 Human sec
41	18	72.0	254	8	ADP31396 Human sec
42	18	72.0	279	8	ADP30846 Human sec
43	18	72.0	279	8	ADP31523 Human sec
44	18	72.0	279	8	ADP30844 Human sec
45	18	72.0	281	4	ADW43729 Corn DBE
46	18	72.0	281	8	ADW43729 Corn DBE
47	18	72.0	286	7	ABO79000 Pseudomon
48	18	72.0	297	8	ADP31192 Human sec
49	18	72.0	312	8	ADP31454 Human sec
50	18	72.0	317	8	ADP31615 Human sec
51	18	72.0	339	8	ADP30519 Human sec
52	18	72.0	346	8	ADP31355 Human sec
53	18	72.0	348	8	ADP31441 Human sec
54	18	72.0	348	8	ADP30985 Human sec
55	18	72.0	348	8	ADP31676 Human sec
56	18	72.0	354	8	ADP30783 Human sec
57	18	72.0	354	8	ADP30779 Human sec
58	18	72.0	354	8	ADP30784 Human sec
59	18	72.0	354	8	ADP30778 Human sec
60	18	72.0	360	8	ADP31209 Human sec
61	18	72.0	363	8	ADP30659 Human sec
62	18	72.0	369	8	ADP30582 Human sec
63	18	72.0	372	8	ADP31492 Human sec
64	18	72.0	382	8	ADP31610 Human sec
65	18	72.0	382	8	ADP31221 Human sec
66	18	72.0	387	8	ADP31368 Human sec
67	18	72.0	396	8	ADP31378 Human sec
68	18	72.0	398	7	ABO71161 Pseudomon
69	18	72.0	405	8	ADP31534 Human sec
70	18	72.0	411	8	ADP30729 Human sec
71	18	72.0	426	8	ADP31495 Human sec
72	18	72.0	442	8	ADP31134 Human sec
73	18	72.0	444	8	ADP31640 Human sec
74	18	72.0	471	8	ADP31083 Human sec
75	18	72.0	474	8	ADP31083 Human sec
76	18	72.0	483	8	ADP30852 Human sec
77	18	72.0	483	8	ADP30853 Human sec
78	18	72.0	484	8	ADP31628 Human sec
79	18	72.0	489	8	ADP31380 Human sec
80	18	72.0	519	8	ADP31190 Human sec
81	18	72.0	525	8	ADP31227 Human sec
82	18	72.0	528	8	ADP30529 Human sec
83	18	72.0	528	8	ADP31525 Human sec
84	18	72.0	533	8	ADP31279 Human sec
85	18	72.0	543	8	ADP30872 Human sec
86	18	72.0	549	8	ADP30855 Human sec
87	18	72.0	552	8	ADP31524 Human sec
88	18	72.0	558	8	ADP31254 Human sec
89	18	72.0	558	8	ADP31255 Human sec
90	18	72.0	592	8	ADP30917 Human sec
91	18	72.0	603	8	ADP31645 Human sec
92	18	72.0	610	8	ADP31264 Human sec
93	18	72.0	612	8	ADP31169 Human sec
94	18	72.0	615	8	ADP31132 Human sec
95	18	72.0	615	8	ADP31361 Human sec
96	18	72.0	615	8	ADP30803 Human sec
97	18	72.0	615	8	ADP31360 Human sec

98	18	72.0	637	8	ADP31397	Human sec	171	18	72.0	1092	8	ADP31358	Human sec
99	18	72.0	639	8	ADP30861	Human sec	172	18	72.0	1107	8	ADP30504	Human sec
100	18	72.0	639	9	ADZ10464	P. gingiv	173	18	72.0	1110	8	ADP31430	Human sec
101	18	72.0	642	8	ADP30748	Human sec	174	18	72.0	1116	8	ADP31692	Human sec
102	18	72.0	660	8	ADP31606	Human sec	175	18	72.0	1128	8	ADP30483	Human sec
103	18	72.0	665	8	ADP30571	Human sec	176	18	72.0	1133	8	ADP30931	Human sec
104	18	72.0	681	8	ADP30868	Human sec	177	18	72.0	1133	8	ADP30928	Human sec
105	18	72.0	681	8	ADP31582	Human sec	178	18	72.0	1133	8	ADP30933	Human sec
106	18	72.0	687	8	ADP31364	Human sec	179	18	72.0	1133	8	ADP30939	Human sec
107	18	72.0	699	8	ADP30747	Human sec	180	18	72.0	1165	8	ADP30879	Human sec
108	18	72.0	705	8	ADP31623	Human sec	181	18	72.0	1167	8	ADP30502	Human sec
109	18	72.0	708	8	ADP31191	Human sec	182	18	72.0	1168	8	ADP31046	Human sec
110	18	72.0	711	8	ADP31440	Human sec	183	18	72.0	1170	8	ADP30919	Human sec
111	18	72.0	729	8	ADP30541	Human sec	184	18	72.0	1170	8	ADP30922	Human sec
112	18	72.0	732	8	ADP30650	Human sec	185	18	72.0	1184	7	ADP06553	Human Ckb
113	18	72.0	735	8	ADP31010	Human sec	186	18	72.0	1189	8	ADP31043	Human sec
114	18	72.0	738	8	ADP31288	Human sec	187	18	72.0	1191	8	ADP30993	Human sec
115	18	72.0	750	8	ADP30857	Human sec	188	18	72.0	1192	8	ADP30964	Human sec
116	18	72.0	750	8	ADP31131	Human sec	189	18	72.0	1199	8	ADP31044	Human sec
117	18	72.0	750	8	ADP30856	Human sec	190	18	72.0	1200	8	ADP31603	Human sec
118	18	72.0	756	8	ADP31039	Human sec	191	18	72.0	1205	8	ADP30935	Human sec
119	18	72.0	762	8	ADP31012	Human sec	192	18	72.0	1215	8	ADP30898	Human sec
120	18	72.0	768	8	ADP30912	Human sec	193	18	72.0	1215	8	ADP30900	Human sec
121	18	72.0	768	8	ADP31362	Human sec	194	18	72.0	1221	8	ADP31293	Human sec
122	18	72.0	768	8	ADP31363	Human sec	195	18	72.0	1224	8	ADP31426	Human sec
123	18	72.0	771	8	ADP31469	Human sec	196	18	72.0	1227	8	ADP31210	Human sec
124	18	72.0	774	8	ADP30885	Human sec	197	18	72.0	1233	8	ADP30523	Human sec
125	18	72.0	780	8	ADP31077	Human sec	198	18	72.0	1237	2	AAR13791	Fruit fly
126	18	72.0	795	8	ADP31332	Human sec	199	18	72.0	1237	6	AAR30117	Human sec
127	18	72.0	816	8	ADP31245	Human sec	200	18	72.0	1260	8	ADP31533	Human sec
128	18	72.0	826	8	ADP31581	Human sec	201	18	72.0	1284	4	ABB71869	D. melano
129	18	72.0	828	8	ADP30878	Human sec	202	18	72.0	1284	5	ABB808376	Human sec
130	18	72.0	834	8	ADP30738	Human sec	203	18	72.0	1302	8	ADP31695	Human sec
131	18	72.0	839	8	ADP31182	Human sec	204	18	72.0	1314	8	ADP31197	Human sec
132	18	72.0	849	8	ADP31184	Human sec	205	18	72.0	1339	8	ADP30732	Human sec
133	18	72.0	861	8	ADP31021	Human sec	206	18	72.0	1341	8	ADP30888	Human sec
134	18	72.0	861	8	ADP31021	Human sec	207	18	72.0	1344	8	ADP31211	Human sec
135	18	72.0	865	8	ADP30566	Human sec	208	18	72.0	1350	8	ADP31138	Human sec
136	18	72.0	885	8	ADP31011	Human sec	209	18	72.0	1358	8	ADP30995	Human sec
137	18	72.0	888	8	ADP31127	Human sec	210	18	72.0	1362	8	ADP31181	Human sec
138	18	72.0	888	8	ADP30971	Human sec	211	18	72.0	1365	8	ADP31035	Human sec
139	18	72.0	889	8	ADP31648	Human sec	212	18	72.0	1371	8	ADP31646	Human sec
140	18	72.0	890	8	ADP31059	Human sec	213	18	72.0	1371	8	ADP30876	Human sec
141	18	72.0	892	8	ADP31578	Human sec	214	18	72.0	1380	8	ADP31566	Human sec
142	18	72.0	906	9	AEB53801	Drosophila	215	18	72.0	1387	8	ADP30946	Human sec
143	18	72.0	921	8	ADP31522	Human sec	216	18	72.0	1417	8	ADP31160	Human sec
144	18	72.0	930	8	ADP31444	Human sec	217	18	72.0	1431	8	ADP31609	Human sec
145	18	72.0	936	8	ADP31597	Human sec	218	18	72.0	1435	6	ABU09754	Human sec
146	18	72.0	939	8	ADP30726	Human sec	219	18	72.0	1435	6	ABU090136	Human sec
147	18	72.0	945	8	ADP31238	Human sec	220	18	72.0	1435	6	ABU96438	Human sec
148	18	72.0	945	8	ADP31237	Human sec	221	18	72.0	1435	6	ABU99047	Human sec
149	18	72.0	948	8	ADP30586	Human sec	222	18	72.0	1435	6	ABU98262	Human sec
150	18	72.0	951	8	ADP30943	Human sec	223	18	72.0	1435	6	ABU91968	Human sec
151	18	72.0	960	8	ADP31471	Human sec	224	18	72.0	1435	6	ABU85272	Human sec
152	18	72.0	960	8	ADP31470	Human sec	225	18	72.0	1435	6	ABO00411	Human sec
153	18	72.0	966	8	ADP30745	Human sec	226	18	72.0	1435	6	ABU88962	Human sec
154	18	72.0	990	8	ADP31553	Human sec	227	18	72.0	1435	6	ABO06458	Human sec
155	18	72.0	1002	8	ADP30866	Human sec	228	18	72.0	1435	6	ABU95518	Human sec
156	18	72.0	1023	8	ADP30994	Human sec	229	18	72.0	1435	6	ABU95208	Human sec
157	18	72.0	1032	8	ADP30874	Human sec	230	18	72.0	1435	6	ABU90756	Human sec
158	18	72.0	1032	8	ADP30863	Human sec	231	18	72.0	1435	6	ABU93918	Human sec
159	18	72.0	1033	8	ADP30984	Human sec	232	18	72.0	1435	6	ABU86192	Human sec
160	18	72.0	1035	8	ADP31552	Human sec	233	18	72.0	1435	6	ABU82047	Human sec
161	18	72.0	1044	8	ADP31517	Human sec	234	18	72.0	1435	6	ABU07908	Human sec
162	18	72.0	1056	8	ADP31082	Human sec	235	18	72.0	1435	6	ABU94228	Human sec
163	18	72.0	1057	8	ADP31592	Human sec	236	18	72.0	1435	6	ABO00101	Human sec
164	18	72.0	1059	8	ADP31042	Human sec	237	18	72.0	1435	6	ABU87112	Human sec
165	18	72.0	1066	8	ADP30563	Human sec	238	18	72.0	1435	6	ABU91353	Human sec
166	18	72.0	1076	6	ABU81145	Human PRO	239	18	72.0	1435	6	ABU90446	Human sec
167	18	72.0	1076	6	ABU68845	Human PRO	240	18	72.0	1435	6	ABU97037	Human sec
168	18	72.0	1082	8	ADP30934	Human sec	241	18	72.0	1435	6	ABO05223	Human sec
169	18	72.0	1086	8	ADP31175	Human sec	242	18	72.0	1464	8	ADP31040	Human sec
170	18	72.0	1090	8	ADQ10187	Human pol	243	18	72.0	1480	8	ADP30557	Human sec

244	18	72.0	1518	8	ADP31532	Adp31532	Human	sec	317	18	72.0	2622	8	ADP31663	Adp31663	Human	sec
245	18	72.0	1550	8	ADP30567	Adp30567	Human	sec	318	18	72.0	2709	8	ADP31526	Adp31526	Human	sec
246	18	72.0	1560	8	ADP31631	Adp31631	Human	sec	319	18	72.0	2790	8	ADP31232	Adp31232	Human	sec
247	18	72.0	1584	8	ADP31405	Adp31405	Human	sec	320	18	72.0	2828	8	ADP30938	Adp30938	Human	sec
248	18	72.0	1598	5	ABR09437	Abbr09437	H. influe		321	18	72.0	2833	8	ADP31299	Adp31299	Human	sec
249	18	72.0	1614	8	ADP31529	Adp31529	Human	sec	322	18	72.0	2835	8	ADP30572	Adp30572	Human	sec
250	18	72.0	1617	8	ADP30660	Adp30660	Human	sec	323	18	72.0	2976	8	ADP30724	Adp30724	Human	sec
251	18	72.0	1647	8	ADP31052	Adp31052	Human	sec	324	18	72.0	2980	9	AEA36049	Maize Sta		
252	18	72.0	1662	8	ADP31419	Adp31419	Human	sec	325	18	72.0	3036	8	ADP31595	Adp31595	Human	sec
253	18	72.0	1662	8	ADP31513	Adp31513	Human	sec	326	18	72.0	3144	8	ADP31544	Adp31544	Human	sec
254	18	72.0	1665	8	ADP31187	Adp31187	Human	sec	327	18	72.0	3201	8	ADP31545	Adp31545	Human	sec
255	18	72.0	1719	8	ADP31137	Adp31137	Human	sec	328	18	72.0	3339	8	ADP31219	Adp31219	Human	sec
256	18	72.0	1725	8	ADP30654	Adp30654	Human	sec	329	18	72.0	3390	8	ADP31148	Adp31148	Human	sec
257	18	72.0	1743	6	ABU88255	Abu88255	Novel	hum	330	18	72.0	3411	8	ADP30667	Adp30667	Human	sec
258	18	72.0	1743	6	ABU90134	Abu90134	Novel	hum	331	18	72.0	3447	8	ADP31112	Adp31112	Human	sec
259	18	72.0	1743	6	ABU96436	Abu96436	Novel	hum	332	18	72.0	3465	8	ADP31334	Adp31334	Human	sec
260	18	72.0	1743	6	ABU99045	Abu99045	Novel	hum	333	18	72.0	3579	8	ADP31098	Adp31098	Human	sec
261	18	72.0	1743	6	ABU98260	Abu98260	Novel	hum	334	18	72.0	4440	6	ABU88256	Abu88256	Novel	hum
262	18	72.0	1743	6	ABU91966	Abu91966	Novel	hum	335	18	72.0	4440	6	ABU90135	Abu90135	Novel	hum
263	18	72.0	1743	6	ABU85270	Abu85270	Novel	hum	336	18	72.0	4440	6	ABU96437	Abu96437	Novel	hum
264	18	72.0	1743	6	ABO00409	Ab000409	Novel	hum	337	18	72.0	4440	6	ABU99046	Abu99046	Novel	hum
265	18	72.0	1743	6	ABU88960	Abu88960	Novel	hum	338	18	72.0	4440	6	ABU98261	Abu98261	Novel	hum
266	18	72.0	1743	6	ABO06456	Ab006456	Novel	hum	339	18	72.0	4440	6	ABU91967	Abu91967	Novel	hum
267	18	72.0	1743	6	ABU95516	Abu95516	Novel	hum	340	18	72.0	4440	6	ABU85271	Abu85271	Novel	hum
268	18	72.0	1743	6	ABU95206	Abu95206	Novel	hum	341	18	72.0	4440	6	ABO00410	Ab000410	Novel	hum
269	18	72.0	1743	6	ABU90754	Abu90754	Novel	hum	342	18	72.0	4440	6	ABU88961	Abu88961	Novel	hum
270	18	72.0	1743	6	ABU93916	Abu93916	Novel	hum	343	18	72.0	4440	6	ABO06457	Ab006457	Novel	hum
271	18	72.0	1743	6	ABU86190	Abu86190	Novel	hum	344	18	72.0	4440	6	ABU95517	Abu95517	Novel	hum
272	18	72.0	1743	6	ABU82045	Abu82045	Novel	hum	345	18	72.0	4440	6	ABU95207	Abu95207	Novel	hum
273	18	72.0	1743	6	ABU07906	Abu07906	Novel	hum	346	18	72.0	4440	6	ABU90755	Abu90755	Novel	hum
274	18	72.0	1743	6	ABU94226	Abu94226	Novel	hum	347	18	72.0	4440	6	ABU93917	Abu93917	Novel	hum
275	18	72.0	1743	6	ABO00099	Ab000099	Novel	hum	348	18	72.0	4440	6	ABU86191	Abu86191	Novel	hum
276	18	72.0	1743	6	ABU87110	Abu87110	Novel	hum	349	18	72.0	4440	6	ABU82046	Abu82046	Novel	hum
277	18	72.0	1743	6	ABU91351	Abu91351	Novel	hum	350	18	72.0	4440	6	ABU07907	Abu07907	Novel	hum
278	18	72.0	1743	6	ABU90444	Abu90444	Novel	hum	351	18	72.0	4440	6	ABU94227	Abu94227	Novel	hum
279	18	72.0	1743	6	ABU97035	Abu97035	Novel	hum	352	18	72.0	4440	6	ABO00100	Ab000100	Novel	hum
280	18	72.0	1743	6	ABO05231	Ab005231	Novel	hum	353	18	72.0	4440	6	ABU87111	Abu87111	Novel	hum
281	18	72.0	1746	8	ADP30992	Adp30992	Human	sec	354	18	72.0	4440	6	ABU91352	Abu91352	Novel	hum
282	18	72.0	1776	9	ADY62715	Ady62715	Human	alp	355	18	72.0	4440	6	ABU90445	Abu90445	Novel	hum
283	18	72.0	1782	8	ADP31391	Adp31391	Human	sec	356	18	72.0	4440	6	ABU97036	Abu97036	Novel	hum
284	18	72.0	1782	8	ADP31270	Adp31270	Human	sec	357	18	72.0	4440	6	ABO05232	Ab005232	Novel	hum
285	18	72.0	1789	8	ADP30962	Adp30962	Human	sec	358	18	72.0	4752	8	ADP30585	Adp30585	Human	sec
286	18	72.0	1793	4	ABR60964	Abbr60964	Drosophi	l	359	18	72.0	4752	8	ADP30651	Adp30651	Human	sec
287	18	72.0	1815	8	ADP31601	Adp31601	Human	sec	360	18	72.0	5304	8	ADP30706	Adp30706	Human	sec
288	18	72.0	1827	8	ADP31170	Adp31170	Human	sec	361	18	72.0	5397	8	ADP31068	Adp31068	Human	sec
289	18	72.0	1833	8	ADP30642	Adp30642	Human	sec	362	18	72.0	5514	8	ADP31186	Adp31186	Human	sec
290	18	72.0	1833	8	ADP30889	Adp30889	Human	sec	363	18	72.0	5514	8	ADP31591	Adp31591	Human	sec
291	18	72.0	1933	8	ADP30902	Adp30902	Human	sec	364	18	72.0	6465	8	ADP30705	Adp30705	Human	sec
292	18	72.0	1956	8	ADP31662	Adp31662	Human	sec	365	18	72.0	7285	6	ABU38280	PAMG21-PA		
293	18	72.0	2001	8	ADP31644	Adp31644	Human	sec	366	18	72.0	7339	6	AAO16358	Human tra		
294	18	72.0	2020	8	ADP31056	Adp31056	Human	sec	367	18	72.0	8976	8	ADP31425	Adp31425	Human	sec
295	18	72.0	2058	8	ADP31630	Adp31630	Human	sec	368	18	72.0	9195	8	ADP31494	Adp31494	Human	sec
296	18	72.0	2065	4	ABR63705	Abbr63705	Drosophi	l	369	18	72.0	10944	8	ADP31311	Adp31311	Human	sec
297	18	72.0	2088	8	ADP31178	Adp31178	Human	sec	370	18	72.0	11328	4	ADP31310	Adp31310	Human	sec
298	18	72.0	2091	8	ADP31088	Adp31088	Human	sec	371	18	72.0	14	4	ABR56666	Human SNP		
299	18	72.0	2127	8	ADP31337	Adp31337	Human	sec	372	17	68.0	23	4	AAE12409	Albunin f		
300	18	72.0	2148	8	ADP30974	Adp30974	Human	sec	373	17	68.0	23	7	ADD68071	Human the		
301	18	72.0	2187	8	ADP30882	Adp30882	Human	sec	374	17	68.0	23	9	AEA39569	Nematode		
302	18	72.0	2304	8	ADP31252	Adp31252	Human	sec	375	17	68.0	23	9	AEA39570	Nematode		
303	18	72.0	2307	8	ADP31394	Adp31394	Human	sec	376	17	68.0	23	9	AEA39571	Nematode		
304	18	72.0	2349	8	ADP30959	Adp30959	Human	sec	377	17	68.0	24	7	ADD95109	PCR prime		
305	18	72.0	2391	8	ADP31366	Adp31366	Human	sec	378	17	68.0	24	8	ADU59677	GBP-4 cto		
306	18	72.0	2418	8	ADP31105	Adp31105	Human	sec	379	17	68.0	25	2	AAV21389	Human HUP		
307	18	72.0	2454	8	ADP30469	Adp30469	Human	sec	380	17	68.0	28	4	AAH33045	Peptide #		
308	18	72.0	2484	8	ADP66630	Adp66630	Human	mis	381	17	68.0	28	4	AAH72815	Human Don		
309	18	72.0	2508	6	ADA15721	Ada15721	C. elegan		382	17	68.0	28	4	AAH60195	Human Bra		
310	18	72.0	2535	8	ADP31146	Adp31146	Human	sec	383	17	68.0	28	4	ABG54519	Human liv		
311	18	72.0	2542	8	ADP31594	Adp31594	Human	sec	384	17	68.0	28	5	ABG42643	Human pep		
312	18	72.0	2544	6	ADA15717	Ada15717	C. elegan		385	17	68.0	29	8	ADP30951	Human sec		
313	18	72.0	2547	8	ADP31665	Adp31665	Human	sec	386	17	68.0	31	3	AAH55811	Human sec		
314	18	72.0	2601	6	ADA15773	Ada15773	C. elegan		387	17	68.0	33	9	AEA34230	Optum pop		
315	18	72.0	2616	8	ADP31253	Adp31253	Human	sec	388	17	68.0	35	3	AAV99909	Peptide e		
316	18	72.0	2616	9	ABR49675	Abbr49675	N. mening		389	17	68.0	45	6	ABR98384	Tumour ce		

390	17	68.0	45	7	ADC84730	Adc84730 MCF-7 bre	463	17	68.0	157	7	ABO81786	AbO81786 Pseudomon
391	17	68.0	48	2	AAy13202	Aay13202 Human sec	464	17	68.0	158	7	ABO75923	ABO75923 Pseudomon
392	17	68.0	48	9	ABN34232	Ae342322 Opium pop	465	17	68.0	161	4	AAm21861	AAm21861 Peptide #
393	17	68.0	49	8	ABO56643	ABO56643 Human gen	466	17	68.0	161	4	ABN44230	ABN44230 Peptide #
394	17	68.0	51	4	AAU14915	Aau14915 Novel Don	467	17	68.0	161	4	AAm38186	AAm38186 Peptide #
395	17	68.0	51	4	AAU50747	AAU50747 Prolionib	468	17	68.0	161	4	ABN27105	ABN27105 Protein #
396	17	68.0	51	6	ABM47266	ABm47266 Prolionib	469	17	68.0	161	4	AAm77967	AAm77967 Human Don
397	17	68.0	52	6	AAU51165	AAu51165 Prolionib	470	17	68.0	161	4	AAm65265	AAm65265 Human bra
398	17	68.0	52	6	ABM47684	ABm47684 Prolionib	471	17	68.0	161	4	ABG59605	ABG59605 Human liv
399	17	68.0	54	6	ABM44785	ABm44785 Prolionib	472	17	68.0	161	5	ABG46980	ABG46980 Human pep
400	17	68.0	54	6	ABM41304	ABm41304 Prolionib	473	17	68.0	164	8	ADX88412	ADX88412 Plant ful
401	17	68.0	58	5	ABP07833	ABp07833 Human ORF	474	17	68.0	164	8	ADY75562	ADY75562 Plant ful
402	17	68.0	61	4	AAm19280	AAm19280 Peptide #	475	17	68.0	165	8	ADP30587	ADP30587 Human sec
403	17	68.0	61	4	ABH38486	ABh38486 Peptide #	476	17	68.0	165	8	ADP31041	ADP31041 Human sec
404	17	68.0	61	4	AAm31981	AAm31981 Peptide #	477	17	68.0	167	9	AEA20070	AEA20070 Novel hum
405	17	68.0	61	4	ABH32648	ABh32648 Protein #	478	17	68.0	168	6	ABJ39095	ABJ39095 Molecule
406	17	68.0	61	4	AAm71647	AAm71647 Human Don	479	17	68.0	168	7	ADM04991	ADM04991 Human pro
407	17	68.0	61	4	AAm59112	AAm59112 Human bra	480	17	68.0	168	8	ADP31099	ADP31099 Human sec
408	17	68.0	61	4	ABG53331	ABg53331 Human liv	481	17	68.0	169	7	ABO83767	ABO83767 Pseudomon
409	17	68.0	61	5	ABG41461	ABg41461 Human pep	482	17	68.0	169	7	AAU64434	AAU64434 Prolionib
410	17	68.0	68	4	AAm24008	AAm24008 Human EST	483	17	68.0	173	6	ABM60953	ABM60953 Human DIT
411	17	68.0	68	4	ADP31694	ADp31694 Human sec	484	17	68.0	173	6	ABR41796	ABR41796 Human DIT
412	17	68.0	74	2	AAW94646	AAw94646 TNF-R ext	485	17	68.0	173	6	ABR41365	ABR41365 Pseudomon
413	17	68.0	74	4	AAH69198	AAh69198 Human TNF	486	17	68.0	173	7	ABO83965	ABO83965 Pseudomon
414	17	68.0	74	7	ADH86259	ADh86259 Enterococ	487	17	68.0	175	4	ABH68003	ABH68003 Drosophil
415	17	68.0	74	7	ABO79231	ABO79231 Pseudomon	488	17	68.0	175	7	ABO80411	ABO80411 Pseudomon
416	17	68.0	78	8	ADH18879	ADh18879 Human cel	489	17	68.0	175	7	ABP75865	ABP75865 Human sec
417	17	68.0	79	6	ABM65345	ABm65345 Prolionib	490	17	68.0	177	6	ADP30755	ADP30755 Human sec
418	17	68.0	79	7	ADP35729	ADp35729 Human hep	491	17	68.0	177	8	ADP30820	ADP30820 Human sec
419	17	68.0	81	7	ABO79675	ABO79675 Pseudomon	492	17	68.0	180	8	ADP30825	ADP30825 Human sec
420	17	68.0	84	3	AAy99920	AAy99920 Peptide e	493	17	68.0	180	8	ADP30825	ADP30825 Human sec
421	17	68.0	84	3	ABG26358	ABg26358 Novel hum	494	17	68.0	180	8	ADP31496	ADP31496 Human sec
422	17	68.0	86	7	ADDS5728	ADd55728 Human hep	495	17	68.0	180	8	ADP30828	ADP30828 Human sec
423	17	68.0	90	8	ADP31655	ADp31655 Human sec	496	17	68.0	180	8	ADP30821	ADP30821 Human sec
424	17	68.0	92	4	AAU39635	AAu39635 Prolionib	497	17	68.0	180	8	ADP30826	ADP30826 Human sec
425	17	68.0	92	6	ABM36154	ABm36154 Prolionib	498	17	68.0	180	8	ADP30827	ADP30827 Pseudomon
426	17	68.0	93	6	AAU47955	AAu47955 Prolionib	499	17	68.0	181	7	ABO77341	ABO77341 Pseudomon
427	17	68.0	93	6	ABM44474	ABm44474 Prolionib	500	17	68.0	184	7	ABM88295	ABM88295 Rice abio
428	17	68.0	93	8	ADP30859	ADp30859 Human sec	501	17	68.0	184	9	AEA20071	AEA20071 Novel hum
429	17	68.0	96	3	ADG27138	ADg27138 Zea maye	502	17	68.0	185	9	ADP31109	ADP31109 Human sec
430	17	68.0	101	4	ABG12358	ABg12358 Novel hum	503	17	68.0	185	8	ADP30641	ADP30641 Human sec
431	17	68.0	103	7	ABO80232	ABO80232 Pseudomon	504	17	68.0	189	8	ADP30641	ADP30641 Human sec
432	17	68.0	107	4	AAU67241	AAu67241 Prolionib	505	17	68.0	191	4	AAU43753	AAU43753 Human sec
433	17	68.0	107	6	ABM63760	ABm63760 Prolionib	506	17	68.0	191	6	ABM40272	ABM40272 Prolionib
434	17	68.0	108	3	AAH18584	AAh18584 Arabidops	507	17	68.0	195	8	ADP30590	ADP30590 Human sec
435	17	68.0	108	3	AAH15332	AAh15332 Arabidops	508	17	68.0	195	8	ADP30590	ADP30590 Human sec
436	17	68.0	109	4	ABH11092	ABh11092 Human sec	509	17	68.0	195	8	ADP30837	ADP30837 Human sec
437	17	68.0	109	8	ADX78125	ADx78125 Plant ful	510	17	68.0	203	7	ABO75097	ABO75097 Pseudomon
438	17	68.0	112	4	AAU45902	AAu45902 Prolionib	511	17	68.0	204	7	ABM90226	ABM90226 Rice abio
439	17	68.0	112	5	ABM42431	ABm42431 Prolionib	512	17	68.0	204	8	ADP31421	ADP31421 Human sec
440	17	68.0	114	5	ABP06472	ABp06472 Human ORF	513	17	68.0	204	8	ADP31422	ADP31422 Human sec
441	17	68.0	115	3	AAH64713	AAh64713 Arabidops	514	17	68.0	204	8	ADP31424	ADP31424 Human sec
442	17	68.0	117	3	AAH38749	AAh38749 Arabidops	515	17	68.0	204	8	ADP30545	ADP30545 Human sec
443	17	68.0	117	3	ADH60202	ADh60202 Secretd	516	17	68.0	204	8	ADP31420	ADP31420 Human sec
444	17	68.0	123	8	ADP31326	ADp31326 Human sec	517	17	68.0	207	7	ADM03987	ADM03987 Human pro
445	17	68.0	128	4	AAU50194	AAu50194 Prolionib	518	17	68.0	207	7	ABO71507	ABO71507 Pseudomon
446	17	68.0	128	6	ABM46713	ABm46713 Prolionib	519	17	68.0	208	7	ABO76683	ABO76683 Pseudomon
447	17	68.0	133	8	ADP31069	ADp31069 Human sec	520	17	68.0	211	6	ABR41776	ABR41776 Human DIT
448	17	68.0	134	4	ABG23610	ABg23610 Novel hum	521	17	68.0	211	7	ADP31155	ADP31155 Human dia
449	17	68.0	134	8	ADRO9969	ADr09969 Human pro	522	17	68.0	213	8	ADP31498	ADP31498 Human sec
450	17	68.0	135	8	ADP30703	ADp30703 Human sec	523	17	68.0	213	8	ADP31499	ADP31499 Human sec
451	17	68.0	136	2	AAy29224	AAy29224 Amino aci	524	17	68.0	214	8	ADY06889	ADY06889 Plant ful
452	17	68.0	138	7	ABO75291	ABO75291 Pseudomon	525	17	68.0	219	8	ADP31171	ADP31171 Human sec
453	17	68.0	138	8	ADP31231	ADp31231 Human sec	526	17	68.0	222	8	ADP31379	ADP31379 Human sec
454	17	68.0	140	3	AAH39333	AAh39333 Human thr	527	17	68.0	225	8	ADP31379	ADP31379 Human sec
455	17	68.0	140	4	ABG13377	ABg13377 Novel hum	528	17	68.0	225	8	ADP31379	ADP31379 Human sec
456	17	68.0	140	7	ADG33298	ADg33298 Human nov	529	17	68.0	230	7	ABO84376	ABO84376 Pseudomon
457	17	68.0	140	7	ABO77148	ABO77148 Pseudomon	530	17	68.0	230	7	ABO84376	ABO84376 Pseudomon
458	17	68.0	151	7	ABO83144	ABO83144 Pseudomon	531	17	68.0	230	9	ADM17135	ADM17135 Human gen
459	17	68.0	152	7	ABO71621	ABO71621 Pseudomon	532	17	68.0	237	8	ADY22944	ADY22944 Plant ful
460	17	68.0	154	7	ABO76866	ABO76866 Pseudomon	533	17	68.0	240	8	AAH06827	AAH06827 Thrombom
461	17	68.0	154	7	ABO67900	ABO67900 Pseudomon	534	17	68.0	240	8	ADP31579	ADP31579 Human sec
462	17	68.0	154	7	ABO74421	ABO74421 Pseudomon	535	17	68.0	240	8	ADY77706	ADY77706 Plant ful

536	17	68.0	242	2	AAR06828	Aar06828	Thrombom	609	17	68.0	361	4	ABB11492	Abb11492	Human neu
537	17	68.0	243	8	ADP30983	Adp30983	Human sec	610	17	68.0	366	8	ADP31106	Adp31106	Human sec
538	17	68.0	243	8	ADP31343	Adp31343	Human sec	611	17	68.0	366	8	ADP31670	Adp31670	Human sec
539	17	68.0	243	9	ADM17137	Adm17137	Eucalyptu	612	17	68.0	367	8	ADP30950	Adp30950	Human sec
540	17	68.0	249	2	AAR06826	Aar06826	Thrombom	613	17	68.0	373	5	ABB47351	Abb47351	Listeria
541	17	68.0	249	2	ABO83160	AbO83160	Pseudomon	614	17	68.0	381	8	ADP30655	Adp30655	Human sec
542	17	68.0	252	8	ADP31485	Adp31485	Human sec	615	17	68.0	382	8	ADP31216	Adp31216	Human sec
543	17	68.0	253	2	AAR06825	Aar06825	Thrombom	616	17	68.0	384	8	ADP30656	Adp30656	Human sec
544	17	68.0	253	2	AAR04241	Aar04241	Thrombin-	617	17	68.0	388	6	ABU39484	Abu39484	Protein e
545	17	68.0	253	2	AAR88514	Aar88514	Thrombom	618	17	68.0	390	8	ADP31584	Adp31584	Human sec
546	17	68.0	253	4	ABB08326	Abb08326	Human thr	619	17	68.0	393	8	ADP31345	Adp31345	Human sec
547	17	68.0	254	3	AAV83938	Aay83938	Human thr	620	17	68.0	395	8	ADP30909	Adp30909	Human sec
548	17	68.0	254	3	AAV83937	Aay83937	Human thr	621	17	68.0	399	7	ADP76549	Adp76549	Pseudomon
549	17	68.0	258	8	ADP30597	Adp30597	Human sec	622	17	68.0	399	8	ADP31313	Adp31313	Human sec
550	17	68.0	262	9	ABE843157	Aeb843157	Zea maye	623	17	68.0	402	7	ABO82656	AbO82656	Pseudomon
551	17	68.0	264	8	ADP30788	Adp30788	Human sec	624	17	68.0	402	8	ADP31414	Adp31414	Human sec
552	17	68.0	264	8	ADP31412	Adp31412	Human sec	625	17	68.0	406	7	ABO69985	AbO69985	Pseudomon
553	17	68.0	264	8	ADP31527	Adp31527	Human sec	626	17	68.0	411	8	ADP31104	Adp31104	Human sec
554	17	68.0	267	8	ADP30822	Adp30822	Human sec	627	17	68.0	411	8	ABM84979	Abm84979	Human dila
555	17	68.0	270	8	ADP30500	Adp30500	Human sec	628	17	68.0	412	4	ABUS3134	Abus3134	Human tra
556	17	68.0	273	8	ADP31236	Adp31236	Human sec	629	17	68.0	414	8	ADP31477	Adp31477	Human sec
557	17	68.0	274	8	ADY22941	Ady22941	Plant ful	630	17	68.0	417	8	ADP31432	Adp31432	Human sec
558	17	68.0	275	2	AAR06832	Aar06832	Thrombom	631	17	68.0	417	9	ADM17788	Adm17788	Pinus rad
559	17	68.0	275	2	AAR94609	Aar94609	Human rec	632	17	68.0	420	8	ADP31349	Adp31349	Human sec
560	17	68.0	276	7	ABO69382	AbO69382	Pseudomon	633	17	68.0	421	7	ABO75404	AbO75404	Pseudomon
561	17	68.0	277	2	AAR06834	Aar06834	Thrombom	634	17	68.0	421	8	ADP31159	Adp31159	Human sec
562	17	68.0	278	8	AAY45578	Aau45578	Propionib	635	17	68.0	421	8	ADS23895	AdS23895	Bacterial
563	17	68.0	278	6	ABM42097	Abm42097	Propionib	636	17	68.0	423	8	ADP30819	Adp30819	Human sec
564	17	68.0	278	8	ADP31134	Adp31134	Human sec	637	17	68.0	423	8	ADP31323	Adp31323	Human sec
565	17	68.0	279	8	ADP31107	Adp31107	Human sec	638	17	68.0	429	8	ADP31100	Adp31100	Human sec
566	17	68.0	279	8	ADP31024	Adp31024	Human sec	639	17	68.0	430	7	ABO69654	AbO69654	Pseudomon
567	17	68.0	288	8	ADP31489	Adp31489	Human sec	640	17	68.0	435	8	ADP31102	Adp31102	Human sec
568	17	68.0	288	8	ADP31453	Adp31453	Human sec	641	17	68.0	437	8	ADP31403	Adp31403	Human sec
569	17	68.0	289	7	ADT58146	Adt58146	Plant pol	642	17	68.0	440	6	ADA48432	Ada48432	Rice prot
570	17	68.0	290	7	ABO70728	AbO70728	Pseudomon	643	17	68.0	448	4	ABG20341	AbG20341	Novel hum
571	17	68.0	291	7	ABO68749	AbO68749	Pseudomon	644	17	68.0	450	8	ADP31085	Adp31085	Human sec
572	17	68.0	293	5	ABB92518	Abb92518	Herbicida	645	17	68.0	453	8	ADP31316	Adp31316	Human sec
573	17	68.0	293	4	ADN74129	Adn74129	Thale cre	646	17	68.0	456	8	ADP31636	Adp31636	Human sec
574	17	68.0	294	4	AAY67478	Aau67478	Propionib	647	17	68.0	461	2	AAR45335	Aar45335	Thrombom
575	17	68.0	294	6	ABM63997	Abm63997	Propionib	648	17	68.0	461	8	ADP31634	Adp31634	Human sec
576	17	68.0	294	8	ADP31473	Adp31473	Human sec	649	17	68.0	462	2	AAR45347	Aar45347	Thrombom
577	17	68.0	297	7	ABO83844	AbO83844	Pseudomon	650	17	68.0	462	2	AAR45349	Aar45349	Thrombom
578	17	68.0	299	6	ABU19784	Abu19784	Protein e	651	17	68.0	462	2	AAR45337	Aar45337	Thrombom
579	17	68.0	300	8	ADP30775	Adp30775	Human sec	652	17	68.0	462	2	AAR45339	Aar45339	Thrombom
580	17	68.0	309	8	ADP30862	Adp30862	Human sec	653	17	68.0	462	2	AAR45343	Aar45343	Thrombom
581	17	68.0	309	8	ADP30873	Adp30873	Human sec	654	17	68.0	462	2	AAR45355	Aar45355	Thrombom
582	17	68.0	310	9	AEA39559	Aea39559	Nematode	655	17	68.0	462	2	AAR45341	Aar45341	Thrombom
583	17	68.0	312	7	ADP30476	Adp30476	Human sec	656	17	68.0	462	2	AAR45342	Aar45342	Thrombom
584	17	68.0	315	7	ADD30795	Add30795	Plant yle	657	17	68.0	462	2	AAR45336	Aar45336	Thrombom
585	17	68.0	315	8	AD141783	Ad141783	Plant tra	658	17	68.0	462	2	AAR45348	Aar45348	Thrombom
586	17	68.0	316	8	ADX94660	Adx94660	Plant ful	659	17	68.0	462	2	AAR45350	Aar45350	Thrombom
587	17	68.0	318	8	ADP31135	Adp31135	Human sec	660	17	68.0	462	2	AAR45345	Aar45345	Thrombom
588	17	68.0	320	8	ADP31607	Adp31607	Human sec	661	17	68.0	462	2	AAR45354	Aar45354	Thrombom
589	17	68.0	320	8	ADP31649	Adp31649	Human sec	662	17	68.0	462	2	AAR45338	Aar45338	Thrombom
590	17	68.0	330	8	ADP31050	Adp31050	Human sec	663	17	68.0	462	2	AAR45353	Aar45353	Thrombom
591	17	68.0	331	1	AAPE2847	Aap2847	Sequence	664	17	68.0	462	2	AAR45346	Aar45346	Thrombom
592	17	68.0	333	8	ADP31442	Adp31442	Human sec	665	17	68.0	462	2	AAR45344	Aar45344	Thrombom
593	17	68.0	338	9	ABM96161	Abm96161	M. xanthu	666	17	68.0	463	2	AAR45340	Aar45340	Thrombom
594	17	68.0	338	8	ADL06030	Adl06030	M. catarr	667	17	68.0	466	5	ABB93873	Abb93873	Herbicida
595	17	68.0	339	8	ADP30892	Adp30892	Human sec	668	17	68.0	471	8	ADP31567	Adp31567	Human sec
596	17	68.0	339	8	ADP30702	Adp30702	Human sec	669	17	68.0	472	8	ADP31222	Adp31222	Human sec
597	17	68.0	340	6	ABM66183	Abm66183	Propionib	670	17	68.0	475	2	AAR22032	Aar22032	Truncated
598	17	68.0	344	6	ABO01324	AbO01324	Human pro	671	17	68.0	476	2	AAR78725	Aar78725	Mature th
599	17	68.0	344	7	ADM26628	Adm26628	Hyperther	672	17	68.0	476	2	AAR86377	Aar86377	Modified
600	17	68.0	344	8	ADUN96020	Adun96020	Human NOV	673	17	68.0	478	8	AAR86376	Aar86376	Modified
601	17	68.0	345	8	ADP31016	Adp31016	Human sec	674	17	68.0	478	8	ADP31007	Adp31007	Human sec
602	17	68.0	345	8	ADP31683	Adp31683	Human sec	675	17	68.0	480	2	AAR22013	Aar22013	Truncated
603	17	68.0	349	6	ABM64594	Abm64594	Propionib	676	17	68.0	480	8	ADP31484	Adp31484	Human sec
604	17	68.0	350	8	ADK57315	Adk57315	Plant ful	677	17	68.0	484	4	ABB63251	Abb63251	Drosophil
605	17	68.0	353	8	ADP31558	Adp31558	Human sec	678	17	68.0	484	6	AAR30102	Aar30102	Drosophil
606	17	68.0	356	4	ABG04360	Abg04360	Novel hum	679	17	68.0	484	9	ABE53795	Aeb53795	Drosophil
607	17	68.0	357	8	ADP30505	Adp30505	Human sec	680	17	68.0	486	2	AAR13877	Aar13877	Thrombin-
608	17	68.0	360	8	ADP31439	Adp31439	Human sec	681	17	68.0	490	5	AAR17531	Aae17531	Human thr

682	17	68.0	490	5	AAE23032	AAe23032	Human	thr	755	17	68.0	564	8	ADP31194	Adp31194	Human	sec
683	17	68.0	494	2	AAE24400	AAe24400	Recombina		756	17	68.0	567	9	ADY70520	Ady70520	Human	bet
684	17	68.0	494	2	AAE78727	AAe78727	Thrombomo		757	17	68.0	564	8	ADP31395	Adp31395	Human	sec
685	17	68.0	494	2	AAE78726	AAe78726	Thrombomo		758	17	68.0	568	4	ABE61334	Abbe1334	Drosophi1	
686	17	68.0	494	3	AAE16401	AAe16401	Novel	bug	759	17	68.0	575	1	ABE82070	Abbe2070	Human	thr
687	17	68.0	497	2	AAE94607	AAe94607	Human	rec	760	17	68.0	575	2	AAE11534	AAe11534	Human	thr
688	17	68.0	497	2	AAE69520	AAe69520	IRFM	prot	761	17	68.0	575	2	AAE14341	AAe14341	Human	thr
689	17	68.0	498	2	AAE84185	AAe84185	Human	der	762	17	68.0	575	2	AAE20639	AAe20639	Human	utr
690	17	68.0	498	2	AAE01600	AAe01600	Thrombomo		763	17	68.0	575	2	AAE22189	AAe22189	Human	thr
691	17	68.0	498	3	AAE67402	AAe67402	Novel	sug	764	17	68.0	575	2	AAE31572	AAe31572	Human	thr
692	17	68.0	500	3	AAE69530	AAe69530	Human	thr	765	17	68.0	575	2	AAE41806	AAe41806	Thrombomo	
693	17	68.0	502	4	ABG03462	ABg03462	Novel	hum	766	17	68.0	575	2	AAE43031	AAe43031	Human	thr
694	17	68.0	510	8	ADP31549	ADp31549	Human	sec	767	17	68.0	575	2	AAE73970	AAe73970	Human	thr
695	17	68.0	510	8	ADP31049	ADp31049	Human	sec	768	17	68.0	575	5	AAE17521	AAe17521	Human	full
696	17	68.0	515	2	AAE10617	AAe10617	Soluble	t	769	17	68.0	575	5	AAE23026	AAe23026	Human	thr
697	17	68.0	516	2	AAE22017	AAe22017	Human	thr	770	17	68.0	575	6	ABG72575	ABg72575	Human	thr
698	17	68.0	516	2	AAE22016	AAe22016	Truncated		771	17	68.0	575	6	ABU52408	ABu52408	Human	GPC
699	17	68.0	516	2	AAE22018	AAe22018	Human	thr	772	17	68.0	575	7	ABM78950	ABm78950	Breast	ca
700	17	68.0	516	2	AAE09347	AAe09347	Human	thr	773	17	68.0	575	7	ADP54760	ADp54760	Human	Pro
701	17	68.0	516	2	AAE09348	AAe09348	Human	thr	774	17	68.0	575	7	ADP48168	ADp48168	Human	Pro
702	17	68.0	516	3	AAE83934	AAe83934	Human	thr	775	17	68.0	575	8	ADL24151	ADl24151	Human	NOV
703	17	68.0	516	3	AAE83935	AAe83935	Human	thr	776	17	68.0	575	8	ADN04006	ADn04006	Antibsort	
704	17	68.0	516	3	AAE69529	AAe69529	Human	thr	777	17	68.0	575	8	ADP12616	ADp12616	Protein	e
705	17	68.0	516	6	ABU08701	ABu08701	Thrombomo		778	17	68.0	575	8	ADP31143	ADp31143	Human	sec
706	17	68.0	516	6	ABU08703	ABu08703	Thrombomo		779	17	68.0	575	8	ADP28681	ADp28681	Human	thr
707	17	68.0	516	6	ADP31081	ADp31081	Human	sec	780	17	68.0	575	8	ADP79506	ADp79506	Human	thr
708	17	68.0	516	7	ADP31271	ADp31271	Human	sol	781	17	68.0	575	8	ADR32191	ADR32191	Human	thr
709	17	68.0	516	7	ADP31275	ADp31275	Human	sol	782	17	68.0	575	8	ADU06321	ADu06321	Novel	bro
710	17	68.0	516	8	ADP30871	ADp30871	Human	sec	783	17	68.0	575	7	ABO79310	ABo79310	Pseudomon	
711	17	68.0	516	8	ADP31418	ADp31418	Human	sec	784	17	68.0	575	8	ADP30877	ADp30877	Human	sec
712	17	68.0	519	6	AAE30101	AAe30101	Drosophi1		785	17	68.0	588	8	ADP30877	ADp30877	Human	sec
713	17	68.0	522	8	ADP31081	ADp31081	Human	sec	786	17	68.0	588	2	AAE35478	AAe35478	Lymphocyt	
714	17	68.0	524	4	AAU07370	AAu07370	G	Protein	787	17	68.0	595	2	AAE24017	AAe24017	Human	sol
715	17	68.0	531	8	ADP31540	ADp31540	Human	sec	788	17	68.0	595	2	AAE30936	AAe30936	Human	CD3
716	17	68.0	531	8	ADP31696	ADp31696	Human	sec	789	17	68.0	595	5	AAE50519	AAe50519	Human	tum
717	17	68.0	532	9	AAE52404	AAe52404	Human	thr	790	17	68.0	595	5	AAU78088	AAu78088	Human	CD3
718	17	68.0	533	1	AAE80641	AAe80641	Sequence		791	17	68.0	595	6	ABP97381	ABp97381	Human	CD3
719	17	68.0	535	1	AAE80641	AAe80641	Sequence		792	17	68.0	595	6	ABU04172	ABu04172	Human	exp
720	17	68.0	544	8	ADY06563	ADy06563	Plant	ful	793	17	68.0	595	6	ABU04178	ABu04178	Human	exp
721	17	68.0	549	8	ADP31639	ADp31639	Human	sec	794	17	68.0	595	6	ABU04176	ABu04176	Human	exp
722	17	68.0	551	8	ADP31639	ADp31639	Human	sec	795	17	68.0	595	6	ABU04177	ABu04177	Human	exp
723	17	68.0	552	8	ADP30540	ADp30540	Human	sec	796	17	68.0	595	6	ABU04175	ABu04175	Human	exp
724	17	68.0	552	8	ADP30875	ADp30875	Human	sec	797	17	68.0	595	6	ABU04174	ABu04174	Human	exp
725	17	68.0	554	5	AAE17522	AAe17522	Human	thr	798	17	68.0	595	6	ABU04169	ABu04169	Human	exp
726	17	68.0	554	5	AAE23027	AAe23027	Human	thr	799	17	68.0	595	6	ABU04169	ABu04169	Human	exp
727	17	68.0	555	8	ADP31168	ADp31168	Human	sec	800	17	68.0	595	6	ABP71420	ABp71420	Human	CoE
728	17	68.0	555	8	ADP31168	ADp31168	Human	sec	801	17	68.0	595	6	ADP19427	ADp19427	Human	CD3
729	17	68.0	555	8	ADP31416	ADp31416	Human	sec	802	17	68.0	595	6	ADP28553	ADp28553	Blinding	d
730	17	68.0	555	8	ADP31417	ADp31417	Human	sec	803	17	68.0	595	7	ADP28553	ADp28553	Human	CD3
731	17	68.0	557	5	AAE17592	AAe17592	Human	thr	804	17	68.0	595	8	ADU24289	ADu24289	Human	CD3
732	17	68.0	557	5	AAE17594	AAe17594	Human	thr	805	17	68.0	595	8	ADU24289	ADu24289	Human	CD3
733	17	68.0	557	5	AAE17524	AAe17524	Human	thr	806	17	68.0	595	8	ADQ6799	ADq6799	Human	CD3
734	17	68.0	557	5	AAE17528	AAe17528	Human	thr	807	17	68.0	595	8	ADQ6799	ADq6799	Human	can
735	17	68.0	557	5	AAE17528	AAe17528	Human	thr	808	17	68.0	595	8	ADQ6799	ADq6799	Human	can
736	17	68.0	557	5	AAE17526	AAe17526	Human	thr	809	17	68.0	595	8	ADQ6799	ADq6799	Human	can
737	17	68.0	557	5	AAE17523	AAe17523	Human	thr	810	17	68.0	595	9	ADY30129	ADy30129	Human	CD3
738	17	68.0	557	5	AAE17596	AAe17596	Human	thr	811	17	68.0	595	9	ADY30129	ADy30129	Human	can
739	17	68.0	557	5	AAE17591	AAe17591	Human	thr	812	17	68.0	600	8	ADP30865	ADp30865	Human	sec
740	17	68.0	557	5	AAE17595	AAe17595	Human	thr	813	17	68.0	600	8	ADP30865	ADp30865	Human	sec
741	17	68.0	557	5	AAE17597	AAe17597	Human	thr	814	17	68.0	600	8	ADP31150	ADp31150	Human	sec
742	17	68.0	557	5	AAE17525	AAe17525	Human	thr	815	17	68.0	605	8	ADP30507	ADp30507	Human	sec
743	17	68.0	557	5	AAE17529	AAe17529	Human	thr	816	17	68.0	605	8	ADP31229	ADp31229	Human	sec
744	17	68.0	557	5	AAE17530	AAe17530	Human	thr	817	17	68.0	609	8	ADP31263	ADp31263	Human	sec
745	17	68.0	557	5	AAE17530	AAe17530	Human	thr	818	17	68.0	609	8	ADP31411	ADp31411	Human	sec
746	17	68.0	557	5	AAE17598	AAe17598	Human	thr	819	17	68.0	612	8	ADP31064	ADp31064	Human	sec
747	17	68.0	557	5	AAE17527	AAe17527	Human	thr	820	17	68.0	617	8	ADN24090	ADn24090	Bacterial	
748	17	68.0	557	5	AAE23031	AAe23031	Human	thr	821	17	68.0	621	8	ADP31147	ADp31147	Human	sec
749	17	68.0	557	5	AAE23029	AAe23029	Human	thr	822	17	68.0	623	9	AAE39561	AAe39561	Nematode	
750	17	68.0	557	5	AAE23030	AAe23030	Human	thr	823	17	68.0	624	8	ADP31324	ADp31324	Human	sec
751	17	68.0	557	8	ADP313028	ADp313028	Human	thr	824	17	68.0	624	8	ADP31325	ADp31325	Human	sec
752	17	68.0	557	8	ADQ65702	ADq65702	Novel	hum	825	17	68.0	627	8	ADP31388	ADp31388	Human	sec
753	17	68.0	558	8	ADP31257	ADp31257	Human	sec	826	17	68.0	629	8	ADP31295	ADp31295	Human	sec
754	17	68.0	561	4	ABE70236	ABe70236	Drosophi1		827	17	68.0	631	4	ABE21715	ABe21715	Novel	hum

828	17	68.0	631	6	ABG72572	Human	thr	Adp31282	Human	sec	901	17	68.0	813	8	ADP31282	Adp31282	Human	sec
829	17	68.0	639	8	ADP31521	Human	sec	Adp30561	Human	sec	902	17	68.0	813	8	ADP30561	Adp30561	Human	sec
830	17	68.0	642	8	ADP31265	Human	sec	Abb68374	Drosophila	sec	903	17	68.0	814	4	ABB68374	Abb68374	Drosophila	sec
831	17	68.0	654	8	ADP30666	Human	sec	Adp30679	Human	sec	904	17	68.0	821	8	ADP30679	Adp30679	Human	sec
832	17	68.0	658	8	ADP31226	Human	sec	Adp30680	Human	sec	905	17	68.0	821	8	ADP30680	Adp30680	Human	sec
833	17	68.0	666	8	ADP31547	Human	sec	Adp31569	Human	sec	906	17	68.0	828	8	ADP31569	Adp31569	Human	sec
834	17	68.0	666	8	ADP30867	Human	sec	Adp31033	Human	sec	907	17	68.0	831	8	ADP31033	Adp31033	Human	sec
835	17	68.0	669	8	ADP31142	Human	sec	Adp31333	Human	sec	908	17	68.0	831	8	ADP31333	Adp31333	Human	sec
836	17	68.0	669	8	ADP31598	Human	sec	Adp31123	Human	sec	909	17	68.0	831	8	ADP31123	Adp31123	Human	sec
837	17	68.0	669	8	ADP31493	Human	sec	Adp31179	Human	sec	910	17	68.0	831	8	ADP31179	Adp31179	Human	sec
838	17	68.0	670	8	ADP31001	Human	sec	Adp30663	Human	sec	911	17	68.0	843	8	ADP30663	Adp30663	Human	sec
839	17	68.0	675	8	ADP31438	Human	sec	ADP30969	Human	sec	912	17	68.0	852	8	ADP30969	ADP30969	Human	sec
840	17	68.0	677	8	ADP31585	Human	sec	ADP30664	Human	sec	913	17	68.0	852	8	ADP30664	ADP30664	Human	sec
841	17	68.0	681	8	ADN05602	Antiposort		ADP30972	Human	sec	914	17	68.0	858	4	ADP30972	ADP30972	Human	sec
842	17	68.0	681	8	ADN05602	Antiposort		Abb60597	Drosophila	sec	915	17	68.0	858	4	ABB60597	Abb60597	Drosophila	sec
843	17	68.0	687	7	ABO77441	Pseudomon		Adp30736	Human	sec	916	17	68.0	868	8	ADP30736	Adp30736	Human	sec
844	17	68.0	690	8	ADP31686	Human	sec	Adp30646	Human	sec	917	17	68.0	870	8	ADP30646	Adp30646	Human	sec
845	17	68.0	690	8	ADP30750	Human	sec	ADP31220	Human	sec	918	17	68.0	876	8	ADP31220	Adp31220	Human	sec
846	17	68.0	702	8	ADP31518	Human	sec	ADP30487	Human	sec	919	17	68.0	882	8	ADP30487	ADP30487	Human	sec
847	17	68.0	711	8	ADP31215	Human	sec	ADP31688	Human	sec	920	17	68.0	882	8	ADP31688	ADP31688	Human	sec
848	17	68.0	711	8	ADP31652	Human	sec	Adp31198	Human	sec	921	17	68.0	885	8	ADP31198	Adp31198	Human	sec
849	17	68.0	714	8	ADP31535	Human	sec	ADP30554	Human	sec	922	17	68.0	887	8	ADP30554	ADP30554	Human	sec
850	17	68.0	714	8	ADP31561	Human	sec	ADP30548	Human	sec	923	17	68.0	887	8	ADP30548	ADP30548	Human	sec
851	17	68.0	720	5	AAO14994	Laminin-5		ADP31230	Human	sec	924	17	68.0	891	8	ADP31230	ADP31230	Human	sec
852	17	68.0	720	5	ADW47628	Laminin-5		ADP30851	Human	sec	925	17	68.0	894	8	ADP30851	ADP30851	Human	sec
853	17	68.0	725	8	ADP31092	Human	sec	ADP31337	Human	sec	926	17	68.0	900	8	ADP31337	ADP31337	Human	sec
854	17	68.0	727	7	ABM68246	Rice	abio	ABU08492	Human	sec	927	17	68.0	908	6	ABU08492	Abu08492	Alpha-hel	
855	17	68.0	728	8	ADP30508	Human	sec	ADP31459	Human	sec	928	17	68.0	918	8	ADP31459	Adp31459	Human	sec
856	17	68.0	731	4	ABG29843	Novel	hum	ADP30546	Human	sec	929	17	68.0	922	8	ADP30546	Adp30546	Human	sec
857	17	68.0	735	8	ADP31520	Human	sec	AAO14246	Human	pre	930	17	68.0	925	5	AAO14246	AAO14246	Human	pre
858	17	68.0	739	4	ABH70356	Drosophila		ABU20097	Protein	e	931	17	68.0	928	6	ABU20097	Abu20097	Protein	e
859	17	68.0	745	8	ADP31196	Human	sec	ADp31140	Human	sec	932	17	68.0	933	8	ADP31140	Adp31140	Human	sec
860	17	68.0	746	8	ADP30982	Human	sec	ADP31568	Human	sec	933	17	68.0	936	8	ADP31568	ADP31568	Human	sec
861	17	68.0	753	8	ADP30988	Human	sec	ADP31542	Human	sec	934	17	68.0	936	8	ADP31542	ADP31542	Human	sec
862	17	68.0	755	8	ADP31559	Human	sec	ADP31541	Human	sec	935	17	68.0	939	8	ADP31541	ADP31541	Human	sec
863	17	68.0	757	8	ADP30925	Human	sec	ADP30937	Human	sec	936	17	68.0	947	8	ADP30937	ADP30937	Human	sec
864	17	68.0	759	8	ADP31141	Human	sec	ADP31167	Human	sec	937	17	68.0	950	8	ADP31167	ADP31167	Human	sec
865	17	68.0	762	8	ADP30891	Human	sec	ADP31528	Human	sec	938	17	68.0	957	8	ADP31528	ADP31528	Human	sec
866	17	68.0	765	8	ADP31149	Human	sec	ADP30547	Human	sec	939	17	68.0	981	8	ADP30547	ADP30547	Human	sec
867	17	68.0	768	2	AAH27684	Human	bet	ADP31057	Human	sec	940	17	68.0	992	8	ADP31057	ADP31057	Human	sec
868	17	68.0	769	5	ABG66351	Human	ova	ADP31538	Human	sec	941	17	68.0	996	8	ADP31538	ADP31538	Human	sec
869	17	68.0	769	6	ABU56733	Lung	canc	ADG39639	Human	pan	942	17	68.0	1002	8	ADG39639	Adg39639	Human	pan
870	17	68.0	769	7	ADH80492	Ovarian	c	ADP30721	Human	sec	943	17	68.0	1008	8	ADP30721	ADP30721	Human	sec
871	17	68.0	769	7	ADN39527	Cancer/an		ADP31296	Human	sec	944	17	68.0	1010	8	ADP31296	Adp31296	Human	sec
872	17	68.0	769	7	ADN39162	Cancer/an		ADP31268	Human	sec	945	17	68.0	1017	8	ADP31268	ADP31268	Human	sec
873	17	68.0	769	7	ADN39597	Cancer/an		ADP30913	Human	sec	946	17	68.0	1030	8	ADP30913	ADP30913	Human	sec
874	17	68.0	769	7	ADH86616	Human	int	ADP30860	Human	sec	947	17	68.0	1038	8	ADP30860	ADP30860	Human	sec
875	17	68.0	771	8	ADP31244	Human	sec	ADP30998	Human	sec	948	17	68.0	1041	8	ADP30998	ADP30998	Human	sec
876	17	68.0	772	8	ADP30936	Human	sec	ADP31642	Human	sec	949	17	68.0	1048	8	ADP31642	ADP31642	Human	sec
877	17	68.0	774	8	ADP30506	Human	sec	ADP31376	Human	sec	950	17	68.0	1050	8	ADP31376	ADP31376	Human	sec
878	17	68.0	774	8	ADP31373	Human	sec	ADP30886	Human	sec	951	17	68.0	1053	8	ADP30886	ADP30886	Human	sec
879	17	68.0	774	8	ADP31225	Human	sec	ABO68762	Pseudomon		952	17	68.0	1064	7	ABO68762	AbO68762	Pseudomon	
880	17	68.0	775	7	ABH85895	Rice	abio	ADG30973	Human	sec	953	17	68.0	1065	8	ADG30973	Adg30973	Human	sec
881	17	68.0	779	8	ADP30915	Human	sec	ADP31347	Human	sec	954	17	68.0	1065	8	ADP31347	ADP31347	Human	sec
882	17	68.0	780	8	ADP30897	Human	sec	ADP31482	Human	sec	955	17	68.0	1065	8	ADP31482	ADP31482	Human	sec
883	17	68.0	780	2	AAH06824	Thrombomo		ADP31287	Human	sec	956	17	68.0	1065	8	ADP31287	ADP31287	Human	sec
884	17	68.0	782	8	ADP30770	Human	sec	ADP30970	Human	sec	957	17	68.0	1070	8	ADP30970	ADP30970	Human	sec
885	17	68.0	782	8	ADP30901	Human	sec	ADP31447	Human	sec	958	17	68.0	1086	8	ADP31447	ADP31447	Human	sec
886	17	68.0	783	8	ADP31284	Human	sec	ADP30574	Human	sec	959	17	68.0	1099	8	ADP30574	ADP30574	Human	sec
887	17	68.0	783	8	ADP31436	Human	sec	ADP31156	Human	sec	960	17	68.0	1104	8	ADP31156	ADP31156	Human	sec
888	17	68.0	783	8	ADP31398	Human	sec	ADP31452	Human	sec	961	17	68.0	1119	8	ADP31452	ADP31452	Human	sec
889	17	68.0	791	4	ABH58225	Drosophila		ADP30920	Human	sec	962	17	68.0	1125	8	ADP30920	ADP30920	Human	sec
890	17	68.0	791	4	ADP3123	PRO	polyp	ADA15725	C. elegans		963	17	68.0	1128	6	ADA15725	AdA15725	C. elegans	
891	17	68.0	799	2	AAH02194	Human	int	ADP30647	Human	sec	964	17	68.0	1134	8	ADP30647	ADP30647	Human	sec
892	17	68.0	799	5	AAU76337	Human	ant	ADP31537	Human	sec	965	17	68.0	1134	8	ADP31537	ADP31537	Human	sec
893	17	68.0	799	7	ADN95792	Human	BBC	ADP30924	Human	sec	966	17	68.0	1134	8	ADP30924	ADP30924	Human	sec
894	17	68.0	799	7	ADP65252	Human	int	ADP31128	Human	sec	967	17	68.0	1140	8	ADP31128	ADP31128	Human	sec
895	17	68.0	799	8	ADL06537	Human	tum	ADP31130	Human	sec	968	17	68.0	1140	8	ADP31130	ADP31130	Human	sec
896	17	68.0	799	8	ADP31261	Human	sec	ADP30677	Human	sec	969	17	68.0	1141	8	ADP30677	ADP30677	Human	sec
897	17	68.0	799	8	ADQ19150	Human	sec	ADP30929	Human	sec	970	17	68.0	1142	8	ADP30929	ADP30929	Human	sec
898	17	68.0	799	9	AEH81278	Human	int	ADP30965	Human	sec	971	17	68.0	1147	8	ADP30965	ADP30965	Human	sec
899	17	68.0	804	8	ADP31635	Human	sec	ADP30906	Human	sec	972	17	68.0	1171	8	ADP30906	ADP30906	Human	sec
900	17	68.0	807	8	ADP31036	Human	sec	ADP30980	Human	sec	973	17	68.0	1171	8	ADP30980	ADP30980	Human	sec

974	17	68.0	1173	8	ADP31155	Adp31155	Human	sec
975	17	68.0	1179	8	ADP30486	Adp30486	Human	sec
976	17	68.0	1179	8	ADP30577	Adp30577	Human	sec
977	17	68.0	1183	8	ADP30536	Adp30536	Human	sec
978	17	68.0	1192	8	ADP31180	Adp31180	Human	sec
979	17	68.0	1194	8	ADP30682	Adp30682	Human	sec
980	17	68.0	1197	8	ADP31034	Adp31034	Human	sec
981	17	68.0	1197	8	ADP31342	Adp31342	Human	sec
982	17	68.0	1200	9	ADP276047	Adp276047	Mouse	pre
983	17	68.0	1222	8	ADP30501	Adp30501	Human	sec
984	17	68.0	1227	8	ADP31602	Adp31602	Human	sec
985	17	68.0	1239	8	ADP31297	Adp31297	Human	sec
986	17	68.0	1248	8	ADP31346	Adp31346	Human	sec
987	17	68.0	1252	8	ADP30678	Adp30678	Human	sec
988	17	68.0	1265	8	ADP31382	Adp31382	Human	sec
989	17	68.0	1269	8	ADP31560	Adp31560	Human	sec
990	17	68.0	1269	8	ADP31381	Adp31381	Human	sec
991	17	68.0	1269	8	ADU23456	Adu23456	Sulfolobu	
992	17	68.0	1282	8	ADP31328	Adp31328	Human	sec
993	17	68.0	1289	8	ADP30675	Adp30675	Human	sec
994	17	68.0	1300	6	ABU88254	Abu88254	Novel	hum
995	17	68.0	1300	6	ABU90133	Abu90133	Novel	hum
996	17	68.0	1300	6	ABU96435	Abu96435	Novel	hum
997	17	68.0	1300	6	ABU99044	Abu99044	Novel	hum
998	17	68.0	1300	6	ABU98259	Abu98259	Novel	hum
999	17	68.0	1300	6	ABU91965	Abu91965	Novel	hum
1000	17	68.0	1300	6	ABU85269	Abu85269	Novel	hum

ALIGNMENTS

RESULT 1
ID ABB41147 standard; peptide; 59 AA.
XX ABB41147;
AC
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #8653 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX
PT gene expression in human foetal liver.
XX
PS Claim 27; SEQ ID NO 33782; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The

CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human foetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at http://wipo.int/pub/published_pct_sequences
XX
SQ Sequence 59 AA;
XX

Query Match 72.0%; Score 18; DB 4; Length 59;
Best Local Similarity 22.2%; Pred. No. 2.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
Db 13 CXXXXXXC 21

RESULT 2
ID AAM34923 standard; protein; 59 AA.
XX AAM34923;
AC
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #8960 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX
PT gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 35192; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP;
CC see AAI13135-AI157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 59 AA;
XX

Query Match 72.0%; Score 18; DB 4; Length 59;
Best Local Similarity 22.2%; Pred. No. 2.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

Db 13 CSSSSTSTC 21

RESULT 3

ID ABB25185 standard; protein; 59 AA.

AC ABB25185;

DT 23-JAN-2002 (first entry)

DE Protein #7184 encoded by probe for measuring heart cell gene expression.

KW Human; gene expression; heart; microarray; vascular system;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

OS Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000666.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human

XX hearts.

XX Claim 15; SEQ ID NO 26955; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for

XX measuring human gene expression in a sample derived from human heart (see

XX ABA2153-ABA41305). The present sequence is a protein encoded by one such

XX probe. The probes may be used for predicting, measuring and displaying

XX gene expression in samples derived from the human heart via microarrays.

XX By measuring gene expression, the probes are useful for predicting,

XX diagnosing, grading, staging, monitoring and prognosing diseases of the

XX human heart and vascular system e.g. cardiovascular disease,

XX hypertension, cardiac arrhythmias and congenital heart disease. Note: The

XX sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 59 AA;

XX Query Match 72.0%; Score 18; DB 4; Length 59;

XX Best Local Similarity 22.2%; Pred. No. 2.8e+02;

XX Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

Db 13 CSSSSTSTC 21

RESULT 4

ID AAM74807 standard; protein; 59 AA.

AC AAM74807;

XX 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 35113.

KW Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukemia; lymphoma; myeloma.

OS Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000668.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human bone marrow.

XX Example 4; SEQ ID NO 35113; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX bone marrow. They can be used to measure gene expression in bone marrow

XX samples, which may enable the improved diagnosis and treatment of cancers

XX such as lymphoma, leukemia and myeloma. The present sequence is a

XX protein encoded by one of the probes of the invention

XX SQ Sequence 59 AA;

XX Query Match 72.0%; Score 18; DB 4; Length 59;

XX Best Local Similarity 22.2%; Pred. No. 2.8e+02;

XX Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

Db 13 CSSSSTSTC 21

RESULT 5

ID AAM62003 standard; protein; 59 AA.

AC AAM62003;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 34108.

KW Human; brain expressed exon; gene expression analysis; probe; microarray;

KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

OS Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000667.

```

XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
XX Example 4; SEQ ID NO 34108; 650bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention
XX
XX Sequence 59 AA;
SQ

```

```

Query Match          72.0%; Score 18; DB 4; Length 59;
Best Local Similarity 22.2%; Pred. No. 2.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 1 CXXXXXXXC 9
   |
DB 13 CSSSTRSTC 21

```

```

RESULT 6
ABG56589
ID ABG56589 standard; peptide; 59 AA.
XX
XX ABG56589;
XX
XX 25-FEB-2003 (first entry)
XX
XX Human liver peptide, SEQ ID No 35237.
XX
XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
XX hypercholesterolaemia; coronary heart disease.
XX
XX Homo sapiens.
XX
XX WO200157273-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000664.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.

```

```

XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
XX Claim 27; SEQ ID NO 35237; 658bp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (1) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 1109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (1) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABG47348-ABG59930 represent human
XX liver single exon encoded peptides of the invention. Note: The sequence
XX information for this patent does not appear in the printed specification
XX but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 59 AA;
SQ

```

```

Query Match          72.0%; Score 18; DB 4; Length 59;
Best Local Similarity 22.2%; Pred. No. 2.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 1 CXXXXXXXC 9
   |
DB 13 CSSSTRSTC 21

```

```

RESULT 7
ABG44598
ID ABG44598 standard; peptide; 59 AA.
XX
XX ABG44598;
XX
XX 19-AUG-2002 (first entry)
XX
XX Human peptide encoded by genome-derived single exon probe SEQ ID 34263.
XX
XX Human; single exon probe; asthma; lung cancer; COPD; IHD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.
XX
XX Homo sapiens.
XX
XX WO200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US000665.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.

```

XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX

PS Claim 27; SEQ ID NO 34263; 634bp; English.

XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridize at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridization of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridization to a single exon
CC microarray having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karsenger syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a peptide/protein encoded by a single exon probe of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

XX Sequence 59 AA;

Query Match 72.0%; Score 18; DB 5; Length 59;

Best Local Similarity 22.2%; Pred. No. 2.8e+02; Mismatches 0; Gaps 0;

Matches 2; Conservative 0; Indels 0; Gaps 0;
QY 1 CXXXXXXC 9
DB 13 CXXXXXXC 21

RESULT 8

AB21190 AAB21190 standard; protein; 108 AA.

XX AAB21190;

XX 08-JUN-2001 (first entry)

XX Exo14 partial protein.

XX Exocytosis pathway protein; Exo protein; antiparkinsonian; antidiabetic;

XX antiallergic; antiallergic; neuroprotective; anticonvulsant;

XX vulnerability; asthma; inflammation; allergy; Chediak-Higashi syndrome; CHS;

XX Alzheimer's disease; Parkinson's disease; Huntington's disease; diabetes;

XX digestion disorder; wound healing disorder; gene therapy.

XX Mus sp.

XX WO200043419-A2.

XX 27-UTL-2000.

XX 20-JAN-2000; 2000WO-US001431.

XX 20-JAN-1999; 99US-0116534P.

XX 26-JAN-1999; 99US-0117274P.

XX 26-JAN-1999; 99US-0117308P.

XX 26-JAN-1999; 99US-0117312P.

XX 01-FEB-1999; 99US-0118177P.

XX 01-FEB-1999; 99US-0118178P.

XX 09-FEB-1999; 99US-0119286P.

XX 11-FEB-1999; 99US-0119759P.

XX 11-FEB-1999; 99US-0119998P.

XX (RIGB-) RIGEL PHARM INC.

XX Luo Y;

XX WPI; 2000-482908/42.

XX N-PSDB; AAA89569.

XX New nucleic acids encoding Exo proteins which are useful in the

XX diagnosis, treatment or prevention of exocytosis-mediated disorders such

XX as asthma, inflammation and allergies.

XX Disclosure; Page 149-150; 305pp; English.

XX The present sequence is a polypeptide which is associated with the
CC exocytosis pathway. CDNA molecules encoding proteins involved in
CC exocytosis have been isolated by yeast one-hybrid and two-hybrid
CC screening. Novel proteins, termed Exo proteins, have been identified that
CC interact with known exocytosis-associated proteins such as GS27, alpha
CC snap, unc18-1, vamps, snap-23, and the rab family of proteins. Exo
CC proteins and their agonists and antagonists are useful in the diagnosis,
CC treatment or prevention of exocytosis-mediated disorders such as asthma,
CC inflammation, allergies, Chediak-Higashi Syndrome (CHS), Alzheimer's
CC disease, Parkinson's disease, Huntington's disease, diabetes, digestion
CC disorders and wound healing disorders. The nucleic acids, antagonists or
CC agonists of Exo proteins are useful in gene therapy. The nucleic acids
CC are also useful for generating transgenic or knock-out animals which can
CC be used in the development and screening of therapeutically useful
CC reagents
XX

XX Sequence 108 AA;

Query Match 72.0%; Score 18; DB 3; Length 108;

Best Local Similarity 22.2%; Pred. No. 3.2e+02; Mismatches 7; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 CXXXXXXC 9
DB 33 CXXXXXXC 41

RESULT 9

AB075321 ABO75321 standard; protein; 134 AA.

XX ABO75321;

XX 29-UTL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #7496.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX

Best Local Similarity 22.2%; Pred. No. 3.4e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
Db 60 CSTSSAASC 68

RESULT 12

AAM93094
ID AAM93094 standard; protein; 140 AA.

AC AAM93094;

DT 06-NOV-2001 (first entry)

DE Human digestive system antigen SEQ ID NO: 2443.

XX Human; digestive system antigen; gene therapy; cancer; appendicitis;
KM ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum.

XX Homo sapiens.

PN WO200155314-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001MO-US001324.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 14-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-022679P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226686P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229333P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 12-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234423P.

PR 21-SEP-2000; 2000US-0234474P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235634P.

PR 27-SEP-2000; 2000US-0235636P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241212P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246539P.

PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.

PR 17-NOV-2000; 2000US-0249210P.

PR 17-NOV-2000; 2000US-0249211P.

PR 17-NOV-2000; 2000US-0249212P.

PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.

PR 17-NOV-2000; 2000US-0249215P.

PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.

PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX
DR MPI: 2001-502630/55.
DR N-PSDB; NAK88867.
XX
PT Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognostizing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX
PS Claim 11; SEQ ID NO 2443; 986bp; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a digestive system antigen of
CC the invention
CC
SQ Sequence 140 AA;
Query Match 72.0%; Score 18; DB 4; Length 140;
Best Local Similarity 22.2%; Pred. No. 3.4e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 1 CXXXXXXXXX 9
Db 104 CTRATSATC 112
RESULT 13
AAU20102
ID AAU20102 standard; protein; 140 AA.
XX
AC AAU20102;
XX
DT 06-DEC-2001 (first entry)
XX
DE Human liver associated polypeptide #133.
XX
KW Liver associated protein; human; mouse; rabbit; goat; horse; cat; dog;
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW antineumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nocitropic; antibacterial; virocidic; fungicidic; cancer;
KW ophthalmological; vulnerrary; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;

KW anti-infertility.
XX
OS Homo sapiens.
XX
PN WO200155355-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001351.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234233P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-023498P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 29-SEP-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-024121P.
PR 20-OCT-2000; 2000US-024121P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251900P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM,
XX
DR WFI, 2001-457728/49.
DR N-PSDB; AAS31813.
XX
PT Isolated nucleic acid molecule encoding a human liver related protein is
PT used in preventing, treating or ameliorating disorders of the liver
PT particularly cancer of the liver.
XX
XX
PS Claim 11; SEQ ID NO 289; 526pp; English.
XX
XX
CC Sequences AAU19970-AAU20115 represent the liver associated polypeptides
CC of the invention. Liver associated polypeptides and their associated
CC polynucleotides are useful in the diagnosis, treatment and prevention of
CC various types of disorders in e.g. humans, mice, rabbits, goats, horses,
CC cats, dogs, chickens or sheep. A pathological condition can be determined
CC by detecting the presence or absence of a mutation in a liver associated
CC polynucleotide. The treatable disorders include autoimmune diseases such
CC as rheumatoid arthritis, hyperproliferative disorders such as neoplasms
CC of the breast or liver, cardiovascular disorders such as cardiac arrest,
CC cerebrovascular disorders such as cerebral ischaemia, nervous system
CC disorders such as Alzheimer's disease, infections caused by bacteria,
CC viruses and fungi, ocular disorders such as corneal infection, endocrine
CC disorders such as Crohn's disease, renal disorders such as
CC glomerulonephritis and respiratory disorders such as asthma and pleurisy.
CC The polypeptides can also be used to aid wound healing, to prevent skin
CC aging due to sunburn, to maintain organs before transplantation, to
CC regenerate tissues and in chemotaxis. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

Query Match 72.0%; Score 18; DB 4; Length 140;
Best Local Similarity 22.2%; Pred. No. 3.4e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
Db 104 CTATSATAC 112

RESULT 14
ABP40963
ID ABP40963 standard; protein; 140 AA.
XX
AC ABP40963;
XX
DT 24-JUL-2002 (first entry)
XX
DE Human liver antigen HFLVG70, SEQ ID NO:289.
XX
KW Human; liver antigen; liver disorder; hepatic disorder; infection;
KW hepatitis; viral; parasitic; bacterial; fungal; inflammatory condition;
KW cirrhosis; granulomatous hepatitis; toxin damage; drug damage;
KW autoimmune disease; Wilson's disease; primary biliary cirrhosis;
KW neoplastic disorder; cancer; tumour; portal hypertension;
KW gastrointestinal disorder; hepatitis; drug screening; gene therapy;
KW chromosome mapping; forensic analysis; antibody preparation;
KW hepatocytic; cytostatic; antiinflammatory; vincristine; antibacterial;
KW fungicide; parasiticide; antidote; immunosuppressive.
XX
XX
OS Homo sapiens.
XX
PN US2002042096-A1.
XX
XX 11-APR-2002.
PD
XX 17-JAN-2001; 2001US-00764887.
PF

XX	31-JAN-2000;	2000US-0179065P.
PR	04-FEB-2000;	2000US-0180628P.
PR	26-JUN-2000;	2000US-0214886P.
PR	07-JUL-2000;	2000US-0216647P.
PR	07-JUL-2000;	2000US-0216880P.
PR	11-JUL-2000;	2000US-0217487P.
PR	11-JUL-2000;	2000US-0217496P.
PR	14-JUL-2000;	2000US-0218290P.
PR	26-JUL-2000;	2000US-0220963P.
PR	26-JUL-2000;	2000US-0220964P.
PR	14-AUG-2000;	2000US-0224518P.
PR	14-AUG-2000;	2000US-0224519P.
PR	14-AUG-2000;	2000US-0225267P.
PR	14-AUG-2000;	2000US-0225268P.
PR	14-AUG-2000;	2000US-0225270P.
PR	14-AUG-2000;	2000US-0225447P.
PR	14-AUG-2000;	2000US-0225757P.
PR	14-AUG-2000;	2000US-0225758P.
PR	32-AUG-2000;	2000US-0226868P.
PR	30-AUG-2000;	2000US-0228924P.
PR	01-SEP-2000;	2000US-0229287P.
PR	01-SEP-2000;	2000US-0229343P.
PR	01-SEP-2000;	2000US-0229344P.
PR	01-SEP-2000;	2000US-0229345P.
PR	05-SEP-2000;	2000US-0229509P.
PR	05-SEP-2000;	2000US-0229513P.
PR	08-SEP-2000;	2000US-0231413P.
PR	21-SEP-2000;	2000US-0234223P.
PR	21-SEP-2000;	2000US-0234274P.
PR	25-SEP-2000;	2000US-0234997P.
PR	27-SEP-2000;	2000US-0235834P.
PR	29-SEP-2000;	2000US-0236327P.
PR	29-SEP-2000;	2000US-0236367P.
PR	29-SEP-2000;	2000US-0236368P.
PR	29-SEP-2000;	2000US-0236369P.
PR	29-SEP-2000;	2000US-0236370P.
PR	02-OCT-2000;	2000US-0236802P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237040P.
PR	02-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239935P.
PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241809P.
PR	01-NOV-2000;	2000US-0244617P.
PR	17-NOV-2000;	2000US-0249299P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251866P.
PR	08-DEC-2000;	2000US-0251866P.
XX	(ROSE/) ROSEN C A.	
PA	(RUBE/) RUBEN S M.	
PA	(BARA/) BARASH S C.	
XX	Rosen CA, Ruben SM, Barash SC;	
PI		
XX		
DR	WPI; 2002-381944/41.	
DR	N-PSDB; AEN90168.	
XX		
PT	New nucleic acid encoding human liver antigens, useful for diagnosis,	
PT	treatment and prevention of e.g. hepatitis and hepatic cancer, also	
PT	related polypeptides and antibodies.	
XX		
PS	Claim 11; SEQ ID NO 289, 181pp; English.	
XX		
CC	The invention relates to 145 novel human liver antigens (ABP40831-	
CC	ABP40975) and to cDNAs encoding them (ABN90036-ABN90180), and also	
CC	encompasses polypeptides 90% identical and polynucleotides 95% identical	
CC	to the sequences of the invention. The invention additionally relates to	
CC	recombinant vectors and host cells comprising human liver antigen	
CC	polynucleotides, antibodies against human liver antigens, and the use of	

CC litholytic, antirheumatic, antiarthritic, neuroprotective, antidiabetic,
CC anticoagulant, thrombolytic, antiarteriosclerotic, cardiant, haemostatic,
CC antiarrhythmic, ophthalmological, antiarteriosclerotic, vasotropic,
CC osteopathic, nootropic, antiparkinsonian, anticonvulsant, neuroleptic,
CC vasotropic, cytostatic and gynaecological activities. The polypeptides
CC and polynucleotides of the invention may be useful for diagnosis,
CC detection, treatment and/or prevention of disorders of the liver such as
CC viral, fungal, bacterial or parasitic infections, cirrhosis, Wilson's
CC disease, gastrointestinal disorders, pancreatic disorders, gallbladder
CC diseases, immune disorders, blood related disorders, hyperproliferative
CC disorders, cardiovascular disorders, respiratory disorders,
CC musculoskeletal system disorders, neurological diseases, endocrine
CC disorders, reproductive system disorders or developmental and inherited
CC disorders. The current sequence is that of the human liver-related
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was obtained electronically from the USPTO web-
CC site.

Query Match 72.0%; Score 18; DB 7; Length 140;
Best Local Similarity 22.2%; Pred. No. 3.4e+02; Mismatches 0; Gaps 0;
Matches 2; Conservative 0; Indels 0;

Qy 1 CXXXXXXC 9
Db 104 CTATSATAC 112

RESULT 16
ID ADP31453 standard; protein; 144 AA.
XX ADP31453;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2220.

XX Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM Cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

XX WO2004035732-A2.

XX PD 29-APR-2004.

XX 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406640P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406646P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

XX 29-AUG-2002; 2002US-0406666P.

XX 17-SEP-2002; 2002US-0410946P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410957P.

XX 17-SEP-2002; 2002US-0410958P.

XX 17-SEP-2002; 2002US-0410959P.

XX 17-SEP-2002; 2002US-0410960P.

XX 17-SEP-2002; 2002US-0410961P.

XX 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486466P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.

PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 3451; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.

XX Sequence 144 AA;

Query Match 72.0%; Score 18; DB 8; Length 144;
Best Local Similarity 22.2%; Pred. No. 3.5e+02; Mismatches 7; Indels 0; Gaps 0;
Matches 2; Conservative 0;

OY 1 CXXXXXXXXC 9
 DB 7 CAATTAATTC 15

RESULT 17
 ABO83270
 ID ABO83270 standard; protein; 145 AA.

XX ABO83270;

XX 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #15445.

KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

XX US651795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

XX N-PSDB; ABD16841.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,

XX useful as molecular targets for diagnostics, prophylaxis and treatment of

XX pathological conditions resulting from bacterial infection.

XX disclosure; SEQ ID NO 32016; 455bp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the

XX polynucleotides encoding them. The sequences are useful in diagnosis and

XX therapy of pathological conditions, as molecular targets for diagnostics,

XX prophylaxis and treatment of pathological conditions resulting from a

XX bacterial infection, for evaluating a compound, such as a polypeptide,

XX for the ability to bind a P. aeruginosa nucleic acid, as components of

XX effective antibacterial targets, as targets for antibacterial drugs,

XX including anti-P. aeruginosa drugs, as templates for recombinant

XX production of P. aeruginosa-derived peptides or polypeptides, as target

XX components for diagnosis and/or treatment of P. aeruginosa-caused

XX infection, and in detection of P. aeruginosa sequences or other sequences

XX of Pseudomonas species using biochip technology. Sequences ABO67826-

XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The

XX sequence data for this patent did not form part of the printed

XX specification but was obtained in electronic format from USPTO at

XX seqdata.uspto.gov/sequence.html

XX Sequence 145 AA;

XX Query Match 72.0%; Score 18; DB 7; Length 145;

XX Best Local Similarity 22.2%; Pred. No. 3.5e+02;

XX Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

XX OY 1 CXXXXXXXXC 9

XX DB 20 CSATASTTC 28

XX RESULT 18

XX ABO79762

XX ID ABO79762 standard; protein; 147 AA.

AC ABO79762;
 XX 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #11937.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

XX US651795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

XX N-PSDB; ABD13333.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,

XX useful as molecular targets for diagnostics, prophylaxis and treatment of

XX pathological conditions resulting from bacterial infection.

XX disclosure; SEQ ID NO 28508; 455bp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the

XX polynucleotides encoding them. The sequences are useful in diagnosis and

XX therapy of pathological conditions, as molecular targets for diagnostics,

XX prophylaxis and treatment of pathological conditions resulting from a

XX bacterial infection, for evaluating a compound, such as a polypeptide,

XX for the ability to bind a P. aeruginosa nucleic acid, as components of

XX effective antibacterial targets, as targets for antibacterial drugs,

XX including anti-P. aeruginosa drugs, as templates for recombinant

XX production of P. aeruginosa-derived peptides or polypeptides, as target

XX components for diagnosis and/or treatment of P. aeruginosa-caused

XX infection, and in detection of P. aeruginosa sequences or other sequences

XX of Pseudomonas species using biochip technology. Sequences ABO67826-

XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The

XX sequence data for this patent did not form part of the printed

XX specification but was obtained in electronic format from USPTO at

XX seqdata.uspto.gov/sequence.html

XX Sequence 147 AA;

XX Query Match 72.0%; Score 18; DB 7; Length 147;

XX Best Local Similarity 22.2%; Pred. No. 3.5e+02;

XX Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

XX OY 1 CXXXXXXXXC 9

XX DB 27 CSTSTTAAC 35

XX RESULT 19

XX ADP31511

XX ID ADP31511 standard; protein; 153 AA.

XX ADP31511;

XX 12-AUG-2004 (first entry).

XX Human secreted protein SEQ ID #2278.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

XX cancer; inflammatory; immune; human secreted protein.

XX

OS Homo sapiens.
XX
XX WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 29-AUG-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467206P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485232P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486460P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Hsieh L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3509; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 153 AA;

Query Match 72.0%; Score 18; DB 8; Length 153;
Best Local Similarity 22.2%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
Db 79 CTAATTTTC 87

RESULT 20
ADN23734
ID ADN23734 standard; protein; 155 AA.

XX ADN23734;
AC
XX 02-DEC-2004 (first entry)
DT
XX
DE Bacterial polypeptide #6387.
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.
PA (HINK) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WI; 2004-061375/06.
DR New recombinant DNA construct comprising a promoter positioned to provide
XX PT for expression of a polynucleotide encoding a polypeptide from a
XX PT microbial source, useful for producing plants with improved properties.
XX PS Claim 1; SEQ ID NO 6387; 122pp; English.
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification
XX of the cell cycle pathway with plant growth regulators, increased rate of
XX homologous recombination, modified seed oil or protein yield and/or
XX content, improved yield by modification of carbohydrate, nitrogen or
XX phosphorus use and/or uptake, by modification of photosynthesis or by
XX providing improved plant growth and development under at least one stress
XX condition, improved lignin production or improved galactomannan
XX production. This sequence represents a bacterial polypeptide used in the
XX scope of the invention. Note: The sequence data for this patent did not
XX form part of the printed specification but was obtained in electronic
XX format from USPTO at seqdata.uspto.gov/sequence.html.
SQ Sequence 155 AA;
Query Match 72.0%; Score 18; DB 8; Length 155;
Best Local Similarity 22.2%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 CXXXXXXC 9
DB 92 CTTASSSSC 100
RESULT 21
ABO83103
XX ID ABO83103 standard; protein; 156 AA.
XX AC ABO83103;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polypeptide #15278.
XX DE Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX KM Pseudomonas aeruginosa.
XX OS Pseudomonas aeruginosa.
XX PN US651795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WI; 2003-615309/58.

DR N-PSDB; ABD16674.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.
XX PS Disclosure; SEQ ID NO 31849; 455pp; English.
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences ABO67826-
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html.
SQ Sequence 156 AA;
Query Match 72.0%; Score 18; DB 7; Length 156;
Best Local Similarity 22.2%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 CXXXXXXC 9
DB 46 CASAAATAC 54
RESULT 22
ADP30759
XX ID ADP30759 standard; protein; 156 AA.
XX AC ADP30759;
XX DT 12-AUG-2004 (first entry)
XX DE Human secreted protein SEQ ID #1526.
XX DE Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX KW cancer; inflammatory; immune; human secreted protein.
XX OS Homo sapiens.
XX PN WO2004035732-A2.
XX PD 29-APR-2004.
XX PF 28-AUG-2003; 2003WO-US026780.
XX PR 29-AUG-2002; 2002US-0406576P.
XX PR 29-AUG-2002; 2002US-0406579P.
XX PR 29-AUG-2002; 2002US-0406585P.
XX PR 29-AUG-2002; 2002US-0406588P.
XX PR 29-AUG-2002; 2002US-0406608P.
XX PR 29-AUG-2002; 2002US-0406611P.
XX PR 29-AUG-2002; 2002US-0406612P.
XX PR 29-AUG-2002; 2002US-0406616P.
XX PR 29-AUG-2002; 2002US-0406640P.
XX PR 29-AUG-2002; 2002US-0406642P.
XX PR 29-AUG-2002; 2002US-0406646P.
XX PR 29-AUG-2002; 2002US-0406653P.
XX PR 29-AUG-2002; 2002US-0406655P.
XX PR 29-AUG-2002; 2002US-0406666P.
XX PR 17-SEP-2002; 2002US-0410946P.
XX PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472430P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476699P.
PR 09-JUN-2003; 2003US-0476611P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486466P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 2757; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present

CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
SQ Sequence 156 AA;
Query Match 72.0%; Score 18; DB 8; Length 156;
Best Local Similarity 22.2%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 CXXXXXXXXX 9
Db 47 CTATTTTAC 55
RESULT 23
ADP31286
ID ADP31286 standard; protein; 165 AA.
XX
AC ADP31286;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2053.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476611P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang WM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 3284; 428bp; English.
XX
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
XX
SQ Sequence 165 AA;
XX
XX
Query Match 72.0%; Score 18; DB 8; Length 165;
Best Local Similarity 22.2%; Pred. No. 3.6e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
QY 1 CXXXXXXC 9
Db 107 CTTTAATC 115
XX
XX
RESULT 24
ID ADP31174 standard; protein; 165 AA.
XX
XX ADP31174;
AC
XX
XX 12-AUG-2004 (first entry)
XX

DE Human secreted protein SEQ ID #1941.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
PD 29-APR-2004.
XX
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410944P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410966P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476611P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.

Db 61 CTSTSSATC 69

RESULT 28

ADP30493

ID ADP30493 standard; protein; 198 AA.

XX

AC ADP30493;

XX

DT 12-AUG-2004 (first entry)

XX

DE Human secreted protein SEQ ID #1260.

XX

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

XX

OS Homo sapiens.

XX

PN WO2004035732-A2.

XX

PD 29-APR-2004.

XX

PF 28-AUG-2003; 2003WO-US026780.

XX

PR 29-AUG-2002; 2002US-0406576P.

XX

PR 29-AUG-2002; 2002US-0406579P.

PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.

PR 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406616P.

PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.

PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.

PR 17-SRP-2002; 2002US-0410946P.

PR 17-SRP-2002; 2002US-0410947P.

PR 17-SRP-2002; 2002US-0410948P.

PR 17-SRP-2002; 2002US-0410949P.

PR 17-SRP-2002; 2002US-0410953P.

PR 17-SRP-2002; 2002US-0410957P.

PR 17-SRP-2002; 2002US-0410958P.

PR 17-SRP-2002; 2002US-0410959P.

PR 17-SRP-2002; 2002US-0410960P.

PR 17-SRP-2002; 2002US-0410961P.

PR 17-SRP-2002; 2002US-0410962P.

PR 17-SRP-2002; 2002US-0411019P.

PR 17-SRP-2002; 2002US-0411022P.

PR 17-SRP-2002; 2002US-0411023P.

PR 17-SRP-2002; 2002US-0411024P.

PR 17-SRP-2002; 2002US-0411032P.

PR 17-SRP-2002; 2002US-0411035P.

PR 17-SRP-2002; 2002US-0411037P.

PR 17-SRP-2002; 2002US-0411041P.

PR 17-SRP-2002; 2002US-0411045P.

PR 17-SRP-2002; 2002US-0411046P.

PR 17-SRP-2002; 2002US-0411048P.

PR 17-SRP-2002; 2002US-0411052P.

PR 17-SRP-2002; 2002US-0411055P.

PR 17-SRP-2002; 2002US-0411073P.

PR 17-SRP-2002; 2002US-0411082P.

PR 17-SRP-2002; 2002US-0411101P.

PR 17-SRP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463709P.

PR 18-APR-2003; 2003US-0463716P.

PR 18-APR-2003; 2003US-0463732P.

PR 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.

PR 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.

PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.

PR 09-JUN-2003; 2003US-0476641P.

PR 08-JUL-2003; 2003US-0485218P.

PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485224P.

PR 08-JUL-2003; 2003US-0485325P.

PR 14-JUL-2003; 2003US-0486446P.

PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486960P.

PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.

PR 08-AUG-2003; 2003US-0493577P.

XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;

PI Halenbeck RF, Huang MM, Kothakota S, Hsiehan L, Lannemann T;

PI Pierce K, Wang Y, Wong JG, Wu G, Zhang H;

XX

DR MPI; 2004-348438/32.

XX

PT New nucleic acid molecule for diagnosing, preventing or treating diseases

PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,

PT genetic, bacterial and viral diseases.

XX

PS Claim 1; SEQ ID NO 2491; 428bp; English.

XX

CC The present invention relates to an isolated nucleic acid molecule

CC encoding a polypeptide which is believed to be cytostatic.

CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and

CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC immune, metabolic, genetic, bacterial and viral diseases. The present

CC sequence represents a human secreted protein. The present sequence is

CC available on WIPWEB and is not in the specification.

XX

SQ Sequence 198 AA;

Query Match 72.0%; Score 18; DB 8; Length 198;

Best Local Similarity 22.2%; Pred. No. 3; 8e+02; 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 7; Indels 0;

QY 1 CXXXXXXXXC 9

Db 91 CATTTTTC 99

RESULT 29

ADP30492

ID ADP30492 standard; protein; 198 AA.

XX

AC ADP30492;

XX

DT 12-AUG-2004 (first entry)

XX

DE Human secreted protein SEQ ID #1259.

XX

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

XX

OS Homo sapiens.

XX

PN WO2004035732-A2.

XX

PD 29-APR-2004.

XX

PF	26-AUG-2003	2003MO-US026780.
XX	29-AUG-2002	2002US-0406576P.
PR	29-AUG-2002	2002US-0406579P.
PR	29-AUG-2002	2002US-0406585P.
PR	29-AUG-2002	2002US-0406588P.
PR	29-AUG-2002	2002US-0406589P.
PR	29-AUG-2002	2002US-0406611P.
PR	29-AUG-2002	2002US-0406612P.
PR	29-AUG-2002	2002US-0406616P.
PR	29-AUG-2002	2002US-0406640P.
PR	29-AUG-2002	2002US-0406642P.
PR	29-AUG-2002	2002US-0406646P.
PR	29-AUG-2002	2002US-0406653P.
PR	29-AUG-2002	2002US-0406655P.
PR	29-AUG-2002	2002US-0406666P.
PR	17-SEP-2002	2002US-0410946P.
PR	17-SEP-2002	2002US-0410947P.
PR	17-SEP-2002	2002US-0410948P.
PR	17-SEP-2002	2002US-0410949P.
PR	17-SEP-2002	2002US-0410961P.
PR	17-SEP-2002	2002US-0410962P.
PR	17-SEP-2002	2002US-0410957P.
PR	17-SEP-2002	2002US-0410957P.
PR	17-SEP-2002	2002US-0411019P.
PR	17-SEP-2002	2002US-0411022P.
PR	17-SEP-2002	2002US-0411023P.
PR	17-SEP-2002	2002US-0411023P.
PR	17-SEP-2002	2002US-0411024P.
PR	17-SEP-2002	2002US-0411032P.
PR	17-SEP-2002	2002US-0411032P.
PR	17-SEP-2002	2002US-0411037P.
PR	17-SEP-2002	2002US-0411037P.
PR	17-SEP-2002	2002US-0411045P.
PR	17-SEP-2002	2002US-0411045P.
PR	17-SEP-2002	2002US-0411046P.
PR	17-SEP-2002	2002US-0411048P.
PR	17-SEP-2002	2002US-0411052P.
PR	17-SEP-2002	2002US-0411055P.
PR	17-SEP-2002	2002US-0411073P.
PR	17-SEP-2002	2002US-0411082P.
PR	17-SEP-2002	2002US-0411101P.
PR	17-SEP-2002	2002US-0411111P.
PR	18-APR-2003	2003US-0463700P.
PR	18-APR-2003	2003US-0463708P.
PR	18-APR-2003	2003US-0463716P.
PR	18-APR-2003	2003US-0463732P.
PR	02-MAY-2003	2003US-0467199P.
PR	02-MAY-2003	2003US-0467201P.
PR	02-MAY-2003	2003US-0467203P.
PR	02-MAY-2003	2003US-0467230P.
PR	19-MAY-2003	2003US-0471306P.
PR	19-MAY-2003	2003US-0471336P.
PR	22-MAY-2003	2003US-0472420P.
PR	22-MAY-2003	2003US-0472430P.
PR	09-JUN-2003	2003US-0476609P.
PR	09-JUN-2003	2003US-0476641P.
PR	08-JUL-2003	2003US-0485218P.
PR	08-JUL-2003	2003US-0485223P.
PR	08-JUL-2003	2003US-0485224P.
PR	08-JUL-2003	2003US-0485235P.
PR	14-JUL-2003	2003US-0486446P.
PR	14-JUL-2003	2003US-0486480P.
PR	15-JUL-2003	2003US-0486891P.
PR	15-JUL-2003	2003US-0486960P.
PR	08-AUG-2003	2003US-0493341P.
PR	08-AUG-2003	2003US-0493370P.
PR	08-AUG-2003	2003US-0493573P.
PR	08-AUG-2003	2003US-0493577P.
XX	(FIVE-) FIVE	PRIME PHARMACEUTICS INC.
XX		

Query Match	Best Local Similarity	Score 18;	DB 8;	Length 198;
Matches	2; Conservative	0; Mismatches	7; Indels	0; Gaps
QY	1 CXXXXXXXXX 9			
DB	91 CATTTTTC 99			
RESULT 30				
ADP30477				
ID	ADP30477 standard; protein; 198 AA.			
XX	ADP30477;			
DT	12-AUG-2004 (first entry)			
XX	Human secreted protein SEQ ID #1244.			
XX	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;			
KM	Cancer; Inflammatory; Immune; human secreted protein.			
XX	Homo sapiens.			
OS				
XX	WO2004035732-A2.			
FN				
XX				
PD	29-APR-2004.			
XX				
PF	28-AUG-2003; 2003WO-US026780.			
XX				
XX	29-AUG-2002; 2002US-0406576P.			
PR	29-AUG-2002; 2002US-0406579P.			
PR	29-AUG-2002; 2002US-0406585P.			
PR	29-AUG-2002; 2002US-0406588P.			
PR	29-AUG-2002; 2002US-0406608P.			
PR	29-AUG-2002; 2002US-0406611P.			
PR	29-AUG-2002; 2002US-0406612P.			
PR	29-AUG-2002; 2002US-0406616P.			
PR	29-AUG-2002; 2002US-0406640P.			
PR	29-AUG-2002; 2002US-0406642P.			
PR	29-AUG-2002; 2002US-0406646P.			
PR	29-AUG-2002; 2002US-0406653P.			
PR	29-AUG-2002; 2002US-0406655P.			
PR	29-AUG-2002; 2002US-0406666P.			
PR	17-SEP-2002; 2002US-0410944P.			
PR	17-SEP-2002; 2002US-0410948P.			
PR	17-SEP-2002; 2002US-0410949P.			
PR	17-SEP-2002; 2002US-0410953P.			
PR	17-SEP-2002; 2002US-0410957P.			

PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0410969P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467206P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471366P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493331P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Heestir K, Beaurang PA, Behrens D,
PI Haldenbeck RF, Huang WM, Kochakota S, Haishan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 2475; 428pp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
XX Sequence 198 AA;

Query Match 72.0%; Score 18; DB 8; Length 198;
Best Local Similarity 22.2%; Pred. No. 3.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 CXXXXXXXXC 9
Db 54 CATATTTC 62
RESULT 31
ADP30481
ID ADP30481 standard; protein; 198 AA.
XX
AC ADP30481;
XX
XX 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1248.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
PD 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411101P.
XX 17-SEP-2002; 2002US-0411111P.

PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisnan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2489; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic.
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
XX Sequence 198 AA;
SQ
Query Match 72.0%; Score 18; DB 8; Length 198;
Best Local Similarity 22.2%; Pred. No. 3.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 CXXXXXXC 9
DB 91 CATTTC 99
RESULT 33
ADP31281
ID ADP31281 standard; protein; 228 AA.
XX
XX ADP31281;
AC
XX
XX 12-AUG-2004 (first entry)
DT
XX
XX Human secreted protein SEQ ID #2048.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
OS
XX
XX WO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PF
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411039P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-047136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-047609P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-048646P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-048690P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisnan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3279; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic.
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.

XX
SQ Sequence 228 AA;

Query Match 72.0%; Score 18; DB 8; Length 228;

Best Local Similarity 22.2%; Pred. No. 3.9e+02; Mismatches 7; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9

DB 183 CTTTAATC 191

RESULT 34

ABO80501 standard; protein; 229 AA.

XX ABO80501;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #12676.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US651795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX MPI: 2003-615309/58.

XX N-PSDB; ABD14072.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 29247; 455bp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences ABO67826-
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html

XX Sequence 229 AA;

Query Match 72.0%; Score 18; DB 7; Length 229;

Best Local Similarity 22.2%; Pred. No. 3.9e+02; Mismatches 7; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9

DB 155 CATTATAC 163

RESULT 35

ABO77555 standard; protein; 233 AA.

XX ABO77555;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #9710.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US651795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX MPI: 2003-615309/58.

XX N-PSDB; ABD11126.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 26301; 455bp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences ABO67826-
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html

XX Sequence 233 AA;

Query Match 72.0%; Score 18; DB 7; Length 233;

Best Local Similarity 22.2%; Pred. No. 3.9e+02; Mismatches 7; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9

DB 43 CSTASAC 51

RESULT 36

```
ADP31468
ID ADP31468 standard; protein; 234 AA.
XX
AC ADP31468;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2235.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX
PR 29-AUG-2002; 2002US-0406579P.
XX
PR 29-AUG-2002; 2002US-0406585P.
XX
PR 29-AUG-2002; 2002US-0406588P.
XX
PR 29-AUG-2002; 2002US-0406608P.
XX
PR 29-AUG-2002; 2002US-0406611P.
XX
PR 29-AUG-2002; 2002US-0406612P.
XX
PR 29-AUG-2002; 2002US-0406616P.
XX
PR 29-AUG-2002; 2002US-0406640P.
XX
PR 29-AUG-2002; 2002US-0406642P.
XX
PR 29-AUG-2002; 2002US-0406646P.
XX
PR 29-AUG-2002; 2002US-0406653P.
XX
PR 29-AUG-2002; 2002US-0406655P.
XX
PR 29-AUG-2002; 2002US-0406666P.
XX
PR 17-SEP-2002; 2002US-0410946P.
XX
PR 17-SEP-2002; 2002US-0410947P.
XX
PR 17-SEP-2002; 2002US-0410948P.
XX
PR 17-SEP-2002; 2002US-0410949P.
XX
PR 17-SEP-2002; 2002US-0410953P.
XX
PR 17-SEP-2002; 2002US-0410957P.
XX
PR 17-SEP-2002; 2002US-0410958P.
XX
PR 17-SEP-2002; 2002US-0410959P.
XX
PR 17-SEP-2002; 2002US-0410960P.
XX
PR 17-SEP-2002; 2002US-0410961P.
XX
PR 17-SEP-2002; 2002US-0410962P.
XX
PR 17-SEP-2002; 2002US-0411019P.
XX
PR 17-SEP-2002; 2002US-0411022P.
XX
PR 17-SEP-2002; 2002US-0411023P.
XX
PR 17-SEP-2002; 2002US-0411024P.
XX
PR 17-SEP-2002; 2002US-0411032P.
XX
PR 17-SEP-2002; 2002US-0411035P.
XX
PR 17-SEP-2002; 2002US-0411037P.
XX
PR 17-SEP-2002; 2002US-0411041P.
XX
PR 17-SEP-2002; 2002US-0411045P.
XX
PR 17-SEP-2002; 2002US-0411046P.
XX
PR 17-SEP-2002; 2002US-0411048P.
XX
PR 17-SEP-2002; 2002US-0411052P.
XX
PR 17-SEP-2002; 2002US-0411055P.
XX
PR 17-SEP-2002; 2002US-0411073P.
XX
PR 17-SEP-2002; 2002US-0411082P.
XX
PR 17-SEP-2002; 2002US-0411101P.
XX
PR 17-SEP-2002; 2002US-0411111P.
XX
PR 18-APR-2003; 2003US-0463700P.
XX
PR 18-APR-2003; 2003US-0463708P.
XX
PR 18-APR-2003; 2003US-0463716P.
XX
PR 18-APR-2003; 2003US-0463732P.
XX
PR 02-MAY-2003; 2003US-0467199P.
XX
PR 02-MAY-2003; 2003US-0467201P.
XX
PR 02-MAY-2003; 2003US-0467203P.
XX
PR 02-MAY-2003; 2003US-0467206P.
XX
PR 19-MAY-2003; 2003US-0471306P.
XX
PR 19-MAY-2003; 2003US-0471336P.
XX
PR 22-MAY-2003; 2003US-0472420P.
XX
PR 22-MAY-2003; 2003US-0472430P.
XX
PR 09-JUN-2003; 2003US-0476609P.
XX
PR 09-JUN-2003; 2003US-0476619P.
XX
PR 08-JUL-2003; 2003US-0485218P.
XX
PR 08-JUL-2003; 2003US-0485223P.
XX
PR 08-JUL-2003; 2003US-0485224P.
XX
PR 08-JUL-2003; 2003US-0485325P.
XX
PR 14-JUL-2003; 2003US-0486446P.
XX
PR 14-JUL-2003; 2003US-0486480P.
XX
PR 15-JUL-2003; 2003US-0486891P.
XX
PR 15-JUL-2003; 2003US-0486960P.
XX
PR 08-AUG-2003; 2003US-0493341P.
XX
PR 08-AUG-2003; 2003US-0493370P.
XX
PR 08-AUG-2003; 2003US-0493573P.
XX
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX William LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
XX PI Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T,
XX PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX PT genetic, bacterial and viral diseases.
XX
XX PS Claim 1; SEQ ID NO 3466; 428bp; English.
XX
XX
XX CC The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cytostatic,
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPOWEB and is not in the specification.
XX
XX SQ Sequence 234 AA;
XX
XX
XX Query Match 72.0%; Score 18; DB 8; Length 234;
XX Best Local Similarity 22.2%; Pred. No. 3.9e+02;
XX Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
XX QY 1 CXXXXXXXC 9
XX
XX Db 213 CTTATTAC 221
XX
XX
XX RESULT 37
XX ADP31616
XX ID ADP31616 standard; protein; 242 AA.
XX
XX AC ADP31616;
XX
XX DT 12-AUG-2004 (first entry)
XX
XX DE Human secreted protein SEQ ID #2363.
XX
XX KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX KM cancer; inflammatory; immune; human secreted protein.
XX
XX OS Homo sapiens.
XX
XX PN WO2004035732-A2.
XX
XX PD 29-APR-2004.
XX
XX PF 28-AUG-2003; 2003WO-US026780.
XX
XX PR 29-AUG-2002; 2002US-0406576P.
XX
XX PR 29-AUG-2002; 2002US-0406579P.
```


PR	29-AUG-2002	2002US-04065855
PR	29-AUG-2002	2002US-0406588P
PR	29-AUG-2002	2002US-0406508P
PR	29-AUG-2002	2002US-0406612P
PR	29-AUG-2002	2002US-0406611P
PR	29-AUG-2002	2002US-0406612P
PR	29-AUG-2002	2002US-0406616P
PR	29-AUG-2002	2002US-0406640P
PR	29-AUG-2002	2002US-0406642P
PR	29-AUG-2002	2002US-0406646P
PR	29-AUG-2002	2002US-0406653P
PR	29-AUG-2002	2002US-0406655P
PR	29-AUG-2002	2002US-0406666P
PR	17-SEP-2002	2002US-0410946P
PR	17-SEP-2002	2002US-0410947P
PR	17-SEP-2002	2002US-0410948P
PR	17-SEP-2002	2002US-0410949P
PR	17-SEP-2002	2002US-0410961P
PR	17-SEP-2002	2002US-0410962P
PR	17-SEP-2002	2002US-0410957P
PR	17-SEP-2002	2002US-0410958P
PR	17-SEP-2002	2002US-0410959P
PR	17-SEP-2002	2002US-0410960P
PR	17-SEP-2002	2002US-0410961P
PR	17-SEP-2002	2002US-0410962P
PR	17-SEP-2002	2002US-0411019P
PR	17-SEP-2002	2002US-0411022P
PR	17-SEP-2002	2002US-0411023P
PR	17-SEP-2002	2002US-0411024P
PR	17-SEP-2002	2002US-0411032P
PR	17-SEP-2002	2002US-0411035P
PR	17-SEP-2002	2002US-0411037P
PR	17-SEP-2002	2002US-0411041P
PR	17-SEP-2002	2002US-0411045P
PR	17-SEP-2002	2002US-0411046P
PR	17-SEP-2002	2002US-0411048P
PR	17-SEP-2002	2002US-0411052P
PR	17-SEP-2002	2002US-0411055P
PR	17-SEP-2002	2002US-0411073P
PR	17-SEP-2002	2002US-0411082P
PR	17-SEP-2002	2002US-0411111P
PR	17-SEP-2002	2002US-0463708P
PR	18-APR-2003	2003US-0463708P
PR	18-APR-2003	2003US-0463716P
PR	18-APR-2003	2003US-0463732P
PR	02-MAY-2003	2003US-0467199P
PR	02-MAY-2003	2003US-0467201P
PR	02-MAY-2003	2003US-0467203P
PR	02-MAY-2003	2003US-0467230P
PR	19-MAY-2003	2003US-0471306P
PR	19-MAY-2003	2003US-0471336P
PR	22-MAY-2003	2003US-0472420P
PR	22-MAY-2003	2003US-0472430P
PR	09-JUN-2003	2003US-0476609P
PR	09-JUN-2003	2003US-0476641P
PR	08-JUL-2003	2003US-0485218P
PR	08-JUL-2003	2003US-0485223P
PR	08-JUL-2003	2003US-0485224P
PR	08-JUL-2003	2003US-0485225P
PR	14-JUL-2003	2003US-0486446P
PR	14-JUL-2003	2003US-0486480P
PR	15-JUL-2003	2003US-0486911P
PR	15-JUL-2003	2003US-0486960P
PR	08-AUG-2003	2003US-0493341P
PR	08-AUG-2003	2003US-0493370P
PR	08-AUG-2003	2003US-0493373P
PR	08-AUG-2003	2003US-0493377P

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D,
PI Hakenbeck RF, Huang MM, Kothakota S, Hashan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGF, Wu G, Zhang H,
XX

DR	WPI, 2004-348438/32.
XX	
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT	genetic, bacterial and viral diseases.
XX	
PS	Claim 1; SEQ ID NO 3614; 428bp; English.
XX	
CC	The present invention relates to an isolated nucleic acid molecule
CC	encoding a polypeptide which is believed to be cytosolic,
CC	anti-inflammatory, immunosuppressive, antibacterial and virucidal. The
CC	composition and methods are useful for diagnosing, preventing and
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,
CC	immune, metabolic, genetic, bacterial and viral diseases. The present
CC	sequence represents a human secreted protein. The present sequence is
CC	available on WIPOMB and is not in the specification.
XX	
SO	Sequence 242 AA;
QY	Query Match 72.0%, Score 18; DB 8; Length 242;
Db	148 CTAATAAAC 156
Matches	2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY	1 CXXXXXXXC 9
Db	148 CTAATAAAC 156
RESULT 38	
AAU75930	
ID	AAU75930 standard; protein, 246 AA.
XX	
AC	AAU75930;
XX	
DT	08-MAY-2002 (first entry)
DE	Androgen receptor ligand binding domain (AR-LBD).
XX	
XX	Androgen; receptor; ligand binding domain; human;
KW	androgen receptor modulator; SARH; hormone-dependent tumour;
KW	prostate cancer; hirsutism; acne; seborrhea; Alzheimer's disease;
KW	androgenic alopecia; hypogonadism; hyperlipidosis; prostate cancer;
KW	benign prostate hypertrophy; adenoma; tumour; pancreatic cancer;
KW	vascular endothelial growth factor; VEGF; antiangiogenic agent;
KW	osteoporosis; spermatogenesis; libido; cachexia; endometriosis;
KW	polycystic ovary syndrome; anorexia; male menopause;
KW	male hormone replacement; female sexual dysfunction; gynaecomastia;
KW	male sexual dysfunction; muscular atrophy; muscle wasting;
KW	protein co-ordinate data.
XX	
OS	Homo sapiens.
XX	
FI	Key
FT	Region
FT	/label= Helix-1
FT	25..49
FT	/label= Helix-3
FT	58..86
FT	/label= Helix4/5
FT	101..105
FT	/label= Helix-6
FT	112..125
FT	/label= Helix-7
FT	130..140
FT	/label= Helix-8
FT	154..169
FT	/label= Helix-9
FT	180..211
FT	/label= Helix-10/11
FT	222..230
FT	/label= Helix-12
XX	
PN	WO200200617-A2.

PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411010P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471366P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486466P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX MPI; 2004-348438/32.
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

XX
XX Claim 1; SEQ ID NO 2752; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic.
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
XX Sequence 249 AA;
SQ
Query Match 72.0%; Score 18; DB 8; Length 249;
Best Local Similarity 22.2%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 CXXXXXXXXC 9
Db 238 CTTAAAAAC 246
RESULT 41
ID ADP31396 standard; protein; 254 AA.
XX
XX ADP31396;
AC
XX 12-AUG-2004 (first entry)
DT
XX
XX Human secreted protein SEQ ID #2163.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
OS
XX Homo sapiens.
OS
XX WO2004035732-A2.
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.

XX	17-SEP-2002; 2002US-0411032P.
PR	17-SEP-2002; 2002US-0411035P.
PR	17-SEP-2002; 2002US-0411037P.
PR	17-SEP-2002; 2002US-0411041P.
PR	17-SEP-2002; 2002US-0411045P.
PR	17-SEP-2002; 2002US-0411046P.
PR	17-SEP-2002; 2002US-0411048P.
PR	17-SEP-2002; 2002US-0411052P.
PR	17-SEP-2002; 2002US-0411055P.
PR	17-SEP-2002; 2002US-0411073P.
PR	17-SEP-2002; 2002US-0411082P.
PR	17-SEP-2002; 2002US-0411101P.
PR	17-SEP-2002; 2002US-0411111P.
PR	18-APR-2003; 2003US-0463700P.
PR	18-APR-2003; 2003US-0463708P.
PR	18-APR-2003; 2003US-0463716P.
PR	18-APR-2003; 2003US-0463732P.
PR	02-MAY-2003; 2003US-0467199P.
PR	02-MAY-2003; 2003US-0467201P.
PR	02-MAY-2003; 2003US-0467203P.
PR	02-MAY-2003; 2003US-0467230P.
PR	19-MAY-2003; 2003US-0471306P.
PR	19-MAY-2003; 2003US-0471336P.
PR	22-MAY-2003; 2003US-0472420P.
PR	22-MAY-2003; 2003US-0472430P.
PR	09-JUN-2003; 2003US-0476609P.
PR	09-JUN-2003; 2003US-0476641P.
PR	08-JUL-2003; 2003US-0485218P.
PR	08-JUL-2003; 2003US-0485223P.
PR	08-JUL-2003; 2003US-0485224P.
PR	08-JUL-2003; 2003US-0485225P.
PR	14-JUL-2003; 2003US-0486446P.
PR	14-JUL-2003; 2003US-0486480P.
PR	15-JUL-2003; 2003US-0486681P.
PR	15-JUL-2003; 2003US-0486690P.
PR	08-AUG-2003; 2003US-0493341P.
PR	08-AUG-2003; 2003US-0493370P.
PR	08-AUG-2003; 2003US-0493573P.
PR	08-AUG-2003; 2003US-0493577P.
XX	
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.
XX	
PI	Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D;
PI	Hallenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI	Pierce K, Wang Y, Wong JCP, Wu G, Zhang H;
XX	
DR	WPI, 2004-348438/32.
XX	
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT	genetic, bacterial and viral diseases.
XX	
PS	Claim 1, SEQ ID NO 3394; 428bp; English.
XX	
CC	The present invention relates to an isolated nucleic acid molecule
CC	encoding a polypeptide which is believed to be cytostatic,
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC	composition and methods are useful for diagnosing, preventing and
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,
CC	immune, metabolic, genetic, bacterial and viral diseases. The present
CC	sequence represents a human secreted protein. The present sequence is
CC	available on WIPOMB and is not in the specification.
XX	
SEQ	Sequence 254 AA;

RESULT 42	
ADP30846	
ID	ADP30846 standard; protein, 279 AA.
XX	
AC	ADP30846;
XX	
DT	12-AUG-2004 (first entry)
XX	
DE	Human secreted protein SEQ ID #1613.
XX	
KW	Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM	cancer; inflammatory; immune; human secreted protein.
XX	
OS	Homo sapiens.
XX	
PN	WO2004035732-A2.
XX	
PD	29-APR-2004.
XX	
PF	28-AUG-2003; 2003MO-US026780.
XX	
PR	29-AUG-2002; 2002US-0406576P.
PR	29-AUG-2002; 2002US-0406579P.
PR	29-AUG-2002; 2002US-0406585P.
PR	29-AUG-2002; 2002US-0406588P.
PR	29-AUG-2002; 2002US-0406608P.
PR	29-AUG-2002; 2002US-0406611P.
PR	29-AUG-2002; 2002US-0406612P.
PR	29-AUG-2002; 2002US-0406616P.
PR	29-AUG-2002; 2002US-0406640P.
PR	29-AUG-2002; 2002US-0406642P.
PR	29-AUG-2002; 2002US-0406646P.
PR	29-AUG-2002; 2002US-0406653P.
PR	29-AUG-2002; 2002US-0406655P.
PR	29-AUG-2002; 2002US-0406666P.
PR	17-SEP-2002; 2002US-0410946P.
PR	17-SEP-2002; 2002US-0410947P.
PR	17-SEP-2002; 2002US-0410948P.
PR	17-SEP-2002; 2002US-0410949P.
PR	17-SEP-2002; 2002US-0410953P.
PR	17-SEP-2002; 2002US-0410957P.
PR	17-SEP-2002; 2002US-0410958P.
PR	17-SEP-2002; 2002US-0410959P.
PR	17-SEP-2002; 2002US-0410960P.
PR	17-SEP-2002; 2002US-0410961P.
PR	17-SEP-2002; 2002US-0410962P.
PR	17-SEP-2002; 2002US-0411019P.
PR	17-SEP-2002; 2002US-0411022P.
PR	17-SEP-2002; 2002US-0411023P.
PR	17-SEP-2002; 2002US-0411024P.
PR	17-SEP-2002; 2002US-0411032P.
PR	17-SEP-2002; 2002US-0411035P.
PR	17-SEP-2002; 2002US-0411037P.
PR	17-SEP-2002; 2002US-0411041P.
PR	17-SEP-2002; 2002US-0411045P.
PR	17-SEP-2002; 2002US-0411046P.
PR	17-SEP-2002; 2002US-0411048P.
PR	17-SEP-2002; 2002US-0411052P.
PR	17-SEP-2002; 2002US-0411055P.
PR	17-SEP-2002; 2002US-0411073P.
PR	17-SEP-2002; 2002US-0411082P.
PR	17-SEP-2002; 2002US-0411101P.
PR	17-SEP-2002; 2002US-0411111P.
PR	18-APR-2003; 2003US-0463700P.
PR	18-APR-2003; 2003US-0463708P.
PR	18-APR-2003; 2003US-0463716P.
PR	18-APR-2003; 2003US-0463732P.
PR	02-MAY-2003; 2003US-0467199P.
PR	02-MAY-2003; 2003US-0467201P.
PR	02-MAY-2003; 2003US-0467203P.
PR	02-MAY-2003; 2003US-0467230P.
PR	19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486460P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2844; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
XX Sequence 279 AA;
SQ
Query Match 72.0%; Score 18; DB 8; Length 279;
Best Local Similarity 22.2%; Pred. No. 4.1e+02; 7; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 CXXXXXXXC 9
DB 247 CTAATTAAC 255
RESULT 43
ADP31523
ID ADP31523 standard; protein; 279 AA.
XX
XX ADP31523;
AC
XX 12-AUG-2004 (first entry)
DT
XX Human secreted protein SEQ ID #2290.
DE
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX Homo sapiens.
OS
XX
XX MO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
XX

PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467206P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486460P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3521; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
XX Sequence 279 AA;
SQ
Query Match 72.0%; Score 18; DB 8; Length 279;
Best Local Similarity 22.2%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 CXXXXXXXC 9
DB 66 CATTAAATC 74
RESULT 44
ADP30844
ID ADP30844 standard; protein; 279 AA.
XX
XX ADP30844;
AC
XX 12-AUG-2004 (first entry)
DT
XX
XX Human secreted protein SEQ ID #1611.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
OS
XX WO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PF
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485233P.
PR 08-JUL-2003; 2003US-0485242P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2842; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
XX Sequence 279 AA;
SQ
Query Match 72.0%; Score 18; DB 8; Length 279;

Best Local Similarity 22.2%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 247 CTATTTAAC 255

RESULT 45

AB48050
ID AB48050 standard; protein; 281 AA.

AC AAB48050;

DT 19-MAR-2001 (first entry)

DE Signal transduction protein.

KW Zea mays; maize; signal transduction protein; phytohormone; ethylene;

XX auxin; cytokinin; gibberellin; immunogen.

OS Zea mays.

PN WO20070059-A2.

PD 23-NOV-2000.

PF 28-APR-2000; 2000WO-US011687.

PR 14-MAY-1999; 99US-0134292P.

PR 08-JUL-1999; 99US-0142996P.

PA (PION-) PIONEER HI-BRED INT INC.

PI HelencJarvis TG;

DR WPI; 2001-031929/04.

DR N-PSDB; AAC84271.

PT New signal transduction nucleic acids and encoded proteins useful for
PT regulating phytohormone expression, including ethylene, auxins,
PT cytokinins and gibberellin, to provide control of plant response to
PT environmental stresses.

PS Claim 13; Page 102; 126pp; English.

XX The invention provides Zea mays signal transduction proteins and encoding
CC nucleotide sequences. The nucleic acids are useful for regulating
CC expression of phytohormones, including ethylene, auxins, cytokinins, and
CC gibberellin, to effect developmental changes in plants and provide
CC control of plant response to environmental stresses. They may also be
CC used as probes or amplification primers in the detection, quantitation or
CC isolation of gene transcripts, for detecting mutations in the gene, for
CC monitoring upregulation of expression or changes in enzyme activity in
CC screening assays of compounds, for detection of any number of allelic
CC variants, or for site-directed mutagenesis in eukaryotic cells. They may
CC further be used for recombinant expression of their encoded polypeptides,
CC as immunogens in the preparation or screening of antibodies, and in sense
CC or antisense suppression of genes in a host cell, tissue or plant. The
CC proteins may be used in assays for enzyme agonists or antagonists, as
CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with the proteins. The present sequence represents a signal transduction
CC protein of the invention

XX Sequence 281 AA;

QY Query Match 72.0%; Score 18; DB 4; Length 281;

Best Local Similarity 22.2%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 11 CSATTTSSC 19

RESULT 46

ADM43729
ID ADM43729 standard; protein; 281 AA.

AC ADM43729;

DT 24-MAR-2005 (first entry)

DE Corn DREB-type transcription factor, seq id 2.

KW Drought resistance; crop improvement; cold tolerance;

XX transcription factor; DREB.

OS Zea mays.

PN CN1472222-A.

PD 04-FEB-2004.

PF 29-JUL-2002; 2002CN-00125372.

PR 29-JUL-2002; 2002CN-00125372.

PA (UYOI) UNIV QINGHUA.

PI Liu Q, Qin F, Zhao J;

DR WPI; 2004-317379/30.

DR N-PSDB; ADM43728.

PT DREB transcription factor of corn and its encoding genes and use.
PS Disclosure; SEQ ID NO 2; 18pp; Chinese.

XX The present invention discloses 2 corn DREB-type transcription factors,
CC madREB1 and madREB2, which are important in conditions of cold and
CC drought. Also disclosed are their coding genes and their application in
CC culture the cold and drought resistant plants. The current sequence
CC represents a DREB-type transcription factor.

XX Sequence 281 AA;

QY Query Match 72.0%; Score 18; DB 8; Length 281;

Best Local Similarity 22.2%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 11 CSATTTSSC 19

RESULT 47

AB079000
ID AB079000 standard; protein; 286 AA.

AC AB079000;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #11175.

KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

PN US6551795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

XX

PS Claim 1; SEQ ID NO 3190; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 297 AA;
XX
Query Match 72.0%; Score 18; DB 8; Length 297;
Best Local Similarity 22.2%; Pred. No. 4.2e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 1 CXXXXXXC 9
Db 39 CAATTTAC 47
XX
RESULT 49
ADP31454
ID ADP31454 standard; protein; 312 AA.
XX
AC ADP31454;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2221.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; Inflammatory; Immune; human secreted protein.
XX
OS Homo sapiens.
XX
XX MO2004035732-AA.
XX
PN 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003MO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX
XX 29-AUG-2002; 2002US-0406579P.
XX
XX 29-AUG-2002; 2002US-0406585P.
XX
XX 29-AUG-2002; 2002US-0406588P.
XX
XX 29-AUG-2002; 2002US-0406608P.
XX
XX 29-AUG-2002; 2002US-0406611P.
XX
XX 29-AUG-2002; 2002US-0406612P.
XX
XX 29-AUG-2002; 2002US-0406616P.
XX
XX 29-AUG-2002; 2002US-0406640P.
XX
XX 29-AUG-2002; 2002US-0406642P.
XX
XX 29-AUG-2002; 2002US-0406646P.
XX
XX 29-AUG-2002; 2002US-0406653P.
XX
XX 29-AUG-2002; 2002US-0406655P.
XX
XX 29-AUG-2002; 2002US-0406666P.
XX
XX 17-SEP-2002; 2002US-0410946P.
XX
XX 17-SEP-2002; 2002US-0410947P.
XX
XX 17-SEP-2002; 2002US-0410948P.
XX
XX 17-SEP-2002; 2002US-0410949P.
XX
XX 17-SEP-2002; 2002US-0410953P.
XX
XX 17-SEP-2002; 2002US-0410957P.
XX
XX 17-SEP-2002; 2002US-0410958P.
XX
XX 17-SEP-2002; 2002US-0410959P.
XX
XX 17-SEP-2002; 2002US-0410960P.
XX
XX 17-SEP-2002; 2002US-0410961P.
XX
XX 17-SEP-2002; 2002US-0410962P.
XX
XX 17-SEP-2002; 2002US-0411019P.
XX
XX 17-SEP-2002; 2002US-0411022P.
XX
XX 17-SEP-2002; 2002US-0411023P.
XX
XX 17-SEP-2002; 2002US-0411024P.
XX
XX 17-SEP-2002; 2002US-0411032P.

PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486860P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Heslir K, Beaurang PA, Behrens D;
XX Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
XX Pierce K, Wang Y, Wong JCP, Wu G, Zhang H;
XX WPI, 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3452; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPOMB and is not in the specification.
XX
SQ Sequence 312 AA;
XX
Query Match 72.0%; Score 18; DB 8; Length 312;
Best Local Similarity 22.2%; Pred. No. 4.2e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 1 CXXXXXXC 9
Db 7 CAATTTAC 15

```
RESULT 50
ADP31615
ID ADP31615 standard; protein; 317 AA.
XX
XX ADP31615;
AC
XX
XX 12-AUG-2004 (first entry)
DT
XX
XX Human secreted protein SEQ ID #2382.
DE
XX
XX Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
OS
XX WO2004035732-A2.
PN
XX
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PP
XX
XX 29-AUG-2002; 2002US-0406576P.
PR
XX 29-AUG-2002; 2002US-0406579P.
PR
XX 29-AUG-2002; 2002US-0406585P.
PR
XX 29-AUG-2002; 2002US-0406588P.
PR
XX 29-AUG-2002; 2002US-0406608P.
PR
XX 29-AUG-2002; 2002US-0406611P.
PR
XX 29-AUG-2002; 2002US-0406612P.
PR
XX 29-AUG-2002; 2002US-0406616P.
PR
XX 29-AUG-2002; 2002US-0406640P.
PR
XX 29-AUG-2002; 2002US-0406642P.
PR
XX 29-AUG-2002; 2002US-0406653P.
PR
XX 29-AUG-2002; 2002US-0406655P.
PR
XX 29-AUG-2002; 2002US-0406666P.
PR
XX 17-SEP-2002; 2002US-0410946P.
PR
XX 17-SEP-2002; 2002US-0410947P.
PR
XX 17-SEP-2002; 2002US-0410948P.
PR
XX 17-SEP-2002; 2002US-0410949P.
PR
XX 17-SEP-2002; 2002US-0410953P.
PR
XX 17-SEP-2002; 2002US-0410957P.
PR
XX 17-SEP-2002; 2002US-0410958P.
PR
XX 17-SEP-2002; 2002US-0410959P.
PR
XX 17-SEP-2002; 2002US-0410960P.
PR
XX 17-SEP-2002; 2002US-0410961P.
PR
XX 17-SEP-2002; 2002US-0410962P.
PR
XX 17-SEP-2002; 2002US-0411019P.
PR
XX 17-SEP-2002; 2002US-0411022P.
PR
XX 17-SEP-2002; 2002US-0411023P.
PR
XX 17-SEP-2002; 2002US-0411024P.
PR
XX 17-SEP-2002; 2002US-0411032P.
PR
XX 17-SEP-2002; 2002US-0411035P.
PR
XX 17-SEP-2002; 2002US-0411037P.
PR
XX 17-SEP-2002; 2002US-0411041P.
PR
XX 17-SEP-2002; 2002US-0411045P.
PR
XX 17-SEP-2002; 2002US-0411046P.
PR
XX 17-SEP-2002; 2002US-0411048P.
PR
XX 17-SEP-2002; 2002US-0411052P.
PR
XX 17-SEP-2002; 2002US-0411055P.
PR
XX 17-SEP-2002; 2002US-0411073P.
PR
XX 17-SEP-2002; 2002US-0411082P.
PR
XX 17-SEP-2002; 2002US-0411101P.
PR
XX 17-SEP-2002; 2002US-0411111P.
PR
XX 18-APR-2003; 2003US-0463700P.
PR
XX 18-APR-2003; 2003US-0463708P.
PR
XX 18-APR-2003; 2003US-0463716P.
PR
XX 18-APR-2003; 2003US-0463732P.
PR
XX 02-MAY-2003; 2003US-0467199P.
PR
XX 02-MAY-2003; 2003US-0467201P.
PR
XX 02-MAY-2003; 2003US-0467203P.
PR
XX 02-MAY-2003; 2003US-0467230P.
PR
XX 19-MAY-2003; 2003US-0471306P.
PR
XX 19-MAY-2003; 2003US-0471336P.
```

```
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA
XX
XX Williams LR, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RP, Huang MM, Kothakota S, Haisan L, Linemann T;
PI Pierce K, Wang Y, Wong JCB, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3613; 428pp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
XX Sequence 317 AA:
SQ
Query Match 72.0%; Score 18; DB 8; Length 317;
Best Local Similarity 22.2%; Pred. No. 4.2e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 CXXXXXXXXC 9
DB 145 CATTATATTC 153
```

Search completed: January 4, 2006, 15:56:15
Job time : 62.5913 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 ; Search time 9.31304 Seconds
(without alignments)
92.983 Million cell updates/sec

Title: US-09-932-322-10
Perfect score: 25
Sequence: 1 CXXXXXXC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 28316 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 28316

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :
1: PIR_80: *
2: PIR1: *
3: PIR2: *
4: PIR3: *
5: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	72.0	155	2 T28945	hypothetical prote
2	18	72.0	470	2 T45962	hypothetical prote
3	18	72.0	1237	2 A34598	ecdysone-induced p
4	18	72.0	1394	2 B34598	ecdysone-induced p
5	18	72.0	1443	2 S05979	steroid hormone re
6	17	68.0	199	2 T33355	hypothetical prote
7	17	68.0	212	2 S63342	hypothetical prote
8	17	68.0	214	2 C70812	probable lipoq prot
9	17	68.0	244	2 H72532	hypothetical prote
10	17	68.0	248	2 T40415	hypothetical prote
11	17	68.0	259	2 H71353	conserved hypotet
12	17	68.0	281	2 AC0661	probable lipoprote
13	17	68.0	293	2 T48975	xyloglucan endo-tr
14	17	68.0	298	2 T01948	hypothetical prote
15	17	68.0	315	2 A85019	probable BHLH DNA-
16	17	68.0	373	2 AB1224	cobalamin biosynth
17	17	68.0	373	2 AD1577	cobalamin biosynth
18	17	68.0	415	2 T43352	nuclear receptor N
19	17	68.0	507	2 T33024	hypothetical prote
20	17	68.0	530	2 T48004	multifunctional am
21	17	68.0	533	2 A42143	skin-1 - Caenorhabd
22	17	68.0	543	2 T07739	probable ferriochel
23	17	68.0	575	2 THHUB	chromomodulin pre
24	17	68.0	582	2 T07953	lectin-like protei
25	17	68.0	582	2 T07952	lectin-like protei
26	17	68.0	584	2 B25682	homeotic protein B
27	17	68.0	589	2 S41011	homeotic protein B
28	17	68.0	595	2 A42086	CD30 antigen precu
29	17	68.0	600	2 D83286	hypothetical prote

30	17	68.0	617	2 T15408	hypothetical prote
31	17	68.0	624	2 A35962	laccase (EC 1.10.3
32	17	68.0	769	2 A41029	integrin beta-8 ch
33	17	68.0	791	2 A46140	diacylglycerol kin
34	17	68.0	796	2 B46140	diacylglycerol kin
35	17	68.0	799	2 A38308	integrin beta-5 ch
36	17	68.0	934	2 T05201	hypothetical prote
37	17	68.0	938	2 T41932	hypothetical prote
38	17	68.0	938	2 T15881	hypothetical prote
39	17	68.0	1295	1 S50820	surface protein ty
40	17	68.0	2704	2 S09118	G surface protein
41	17	68.0	2718	2 A23475	G surface protein
42	17	68.0	8243	2 T31307	type I fatty acid
43	16	64.0	52	2 G82010	probable lipoprote
44	16	64.0	87	2 H86753	prophage p12 prote
45	16	64.0	91	2 J50036	Clara cell 10K pro
46	16	64.0	96	1 UGMS	uteroglobin precu
47	16	64.0	96	2 A36581	polychlorinated bi
48	16	64.0	98	2 A10334	conserved hypotet
49	16	64.0	103	2 S76421	ferredoxin [2Fe-2S
50	16	64.0	104	2 J00863	hypothetical 11.6K
51	16	64.0	116	2 S09845	hypothetical prote
52	16	64.0	121	2 A44544	hypothetical prote
53	16	64.0	128	2 S74085	amine dehydrogenas
54	16	64.0	138	2 T21792	lutropin beta chal
55	16	64.0	148	2 B72513	hypothetical prote
56	16	64.0	152	2 H97405	hypothetical prote
57	16	64.0	153	2 F75478	conserved hypotet
58	16	64.0	156	2 T31839	hypothetical prote
59	16	64.0	158	2 A61091	lutropin beta chal
60	16	64.0	159	2 I51373	lutropinizing hormo
61	16	64.0	160	1 S17194	endothelin 2 precu
62	16	64.0	166	2 I51242	lutropinizing hormo
63	16	64.0	166	2 S69590	hypothetical prote
64	16	64.0	182	2 T01909	transcription init
65	16	64.0	190	2 S48101	xyloglucan endo-1,
66	16	64.0	195	2 T19151	hypothetical prote
67	16	64.0	207	2 A69446	conserved hypotet
68	16	64.0	209	2 S51480	drought-induced pr
69	16	64.0	238	2 T04166	thumatin-like pro
70	16	64.0	246	2 A63519	enterobactin synth
71	16	64.0	246	2 T20341	hypothetical prote
72	16	64.0	248	2 T21786	hypothetical prote
73	16	64.0	251	2 A55523	hypothetical prote
74	16	64.0	252	2 T04739	hypothetical prote
75	16	64.0	255	2 T18426	lymphocyte activat
76	16	64.0	260	2 T23033	hypothetical prote
77	16	64.0	261	2 C64049	hypothetical prote
78	16	64.0	261	2 S63604	homobox protein G
79	16	64.0	262	2 S01914	El membrane glycop
80	16	64.0	265	2 H97280	hydroxyethylthiaz
81	16	64.0	266	2 T02807	arsenate reductase
82	16	64.0	266	2 H84409	thiamin biosynthes
83	16	64.0	268	2 S19184	hypothetical prote
84	16	64.0	268	2 H81806	histidine-binding
85	16	64.0	268	2 B81063	amino acid ABC tra
86	16	64.0	269	2 T35647	phosphomethylpyrim
87	16	64.0	275	2 T32813	hypothetical prote
88	16	64.0	276	2 T01177	hypothetical prote
89	16	64.0	276	2 T52349	F-box protein FBX2
90	16	64.0	277	2 A46241	interferon respons
91	16	64.0	278	2 T39749	hypothetical prote
92	16	64.0	279	2 T51017	hypothetical prote
93	16	64.0	286	2 G97442	hypothetical prote
94	16	64.0	286	2 H75551	dihydropterolate sy
95	16	64.0	293	2 T31840	hypothetical prote
96	16	64.0	295	2 S48102	xyloglucan endo-1,
97	16	64.0	299	2 F84785	probable transcrip
98	16	64.0	299	2 F83090	hypothetical prote
99	16	64.0	303	2 T29321	hypothetical prote
100	16	64.0	327	2 A46484	apoptosis-mediatin
101	16	64.0	331	2 A71338	conserved hypotet
102	16	64.0	332	2 T26339	hypothetical prote

979 14 56.0 47 2 S22828
980 14 56.0 47 2 I48849
981 14 56.0 49 2 S00228
982 14 56.0 49 2 S02007
983 14 56.0 50 1 HSPG
984 14 56.0 50 1 HSPG
985 14 56.0 50 2 S22582
986 14 56.0 50 2 S21672
987 14 56.0 50 2 T38209
988 14 56.0 51 1 HSBOS
989 14 56.0 51 1 HSBOS
990 14 56.0 52 2 S03997
991 14 56.0 52 2 S65712
992 14 56.0 53 2 S24596
993 14 56.0 53 2 S75905
994 14 56.0 55 1 B7A58
995 14 56.0 56 1 NIBSSA
996 14 56.0 56 2 T12783
997 14 56.0 57 2 S12957
998 14 56.0 58 1 SMK015
999 14 56.0 58 2 G97810
1000 14 56.0 59 1 FEDVIV

ALIGNMENTS

RESULT 1
T28945
hypothetical protein F07C4.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C/Accession: T28945
R/Miller, N.; Steilyes, L.
submitted to the EMBL Data Library, January 1997
A/Description: The sequence of C. elegans cosmid F07C4.
A/Reference number: Z20546
A/Accession: T28945
A/Status: preliminary; translated from GB/EMBL/DBD
A/Molecule type: DNA
A/Residues: 1-155 <MIL>
A/Cross-references: UNIPROT:P91214; UNIPARC:UPI000007DB2D; EMBL:U80023; PIDN:AAC48012.1
A/Experimental source: strain Bristol N2; clone F07C4
C/Genetics:
A/Gene: CSP.F07C4.2
A/Map position: 5
A/Intons: 116/1
C/Superfamily: tetranectin; C-type lectin homology

Query Match 72.0%; Score 18; DB 2; Length 155;
Best Local Similarity 22.2%; Pred. No. 17;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 92 CTTASSSC 100

RESULT 2
T45962
hypothetical protein F7J8.200 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C/Accession: T45962
R/Bevan, M.; Zimmermann, W.; Gueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Lem
submitted to the Protein Sequence Database, January 2000
A/Reference number: Z23018
A/Accession: T45962
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-470 <BRV>
A/Cross-references: UNIPROT:Q9LFB4; UNIPARC:UPI00000A7D3B; EMBL:AL137189
A/Experimental source: cultivar Columbia; BAC clone F7J8

C/Genetics:
A/Map position: 5
A/Intons: 116/2; 138/3; 162/2; 201/3; 226/2; 250/1; 275/1; 288/2; 307/2; 379/3
A/Note: F7J8.200

Query Match 72.0%; Score 18; DB 2; Length 470;
Best Local Similarity 22.2%; Pred. No. 23;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 23 CSSSSATSC 31

RESULT 3
A34598
ecdysone-induced protein E75A - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 29-Jun-1990 #sequence_revision 29-Jun-1990 #text_change 31-Dec-2004
C/Accession: A34598
R/Segraves, W.A.; Hogness, D.S.
Genes Dev 4, 204-219, 1990
A/Title: The E75 ecdysone-inducible gene responsible for the 75B early puff in Drosophila
A/Reference number: A34598; M01D:90249727; PMID:2110921
A/Accession: A34598
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1237 <SEG>
A/Cross-references: UNIPROT:P17671; UNIPARC:UPI0000129B01; GB:X51548; NID:g7910; PIDN:CA
C/Genetics:
A/Gene: FlyBase:Bip75B
A/Cross-references: FlyBase:FBN0000568
C/Keywords: alternative splicing; DNA binding; nucleus; transcription regulation; zinc f
P;243-520/Domain: erba transforming protein homology <ERBA>

Query Match 72.0%; Score 18; DB 2; Length 1237;
Best Local Similarity 22.2%; Pred. No. 30;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 1070 CSSSSASSC 1078

RESULT 4
B34598
ecdysone-induced protein E75B - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 29-Jun-1990 #sequence_revision 29-Jun-1990 #text_change 31-Dec-2004
C/Accession: B34598
R/Segraves, W.A.; Hogness, D.S.
Genes Dev 4, 204-219, 1990
A/Title: The E75 ecdysone-inducible gene responsible for the 75B early puff in Drosophila
A/Reference number: A34598; M01D:90249727; PMID:2110921
A/Accession: B34598
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1394 <SEG>
A/Cross-references: UNIPROT:P17672; UNIPARC:UPI0000129B01; GB:X51549; NID:g7912; PIDN:CA
C/Genetics:
A/Gene: FlyBase:Bip75B
A/Cross-references: FlyBase:FBN0000568
C/Keywords: alternative splicing; DNA binding; nucleus; transcription regulation; zinc f
P;401-677/Domain: erba transforming protein homology <ERBA>

Query Match 72.0%; Score 18; DB 2; Length 1394;
Best Local Similarity 22.2%; Pred. No. 31;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 1227 CSSSSASSC 1235

RESULT 5

S05979
steroid hormone receptor homolog - fruit fly (Drosophila melanogaster)
N:Alternate names: puff 75B protein
C:Species: Drosophila melanogaster
C>Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 05-Oct-2004
C:Accession: S05979
R:Feigl, G.; Gram, M.; Pongs, O.
Nucleic Acids Res. 17, 7167-7178, 1989
A:Title: A member of the steroid hormone receptor gene family is expressed in the 20-OH-
A:Reference number: S05979; M01D:90016778; PMID:2508058
A:Accession: S05979
A:Molecule type: mRNA
A:Residues: 1-143 <FEI>
A:Cross-references: UNIPROT:P13055; UNIPARC:UPI0000129B02; EMBL:X15586; NID:G7516; PIDN:
C:Genetics:
A:Gene: FlyBase:FlyBase:FBgn0000568
A:Cross-references: FlyBase:FBgn0000568
A:Map position: 3 75B
C:Keywords: alternative splicing; DNA binding; steroid hormone receptor; transcription
F:452-727/Domain: erba transforming protein homology <ERBA>

Query Match 72.0%; Score 18; DB 2; Length 1443;
Best Local Similarity 22.2%; Pred. No. 32;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 1276 CSSSASASC 1284

RESULT 6

hypothetical protein F16G10.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T13355
R:Gutting, S.; Scheet, P.
Submitted to the EMBL Data Library, July 1998
A:Description: The sequence of C. elegans cosmid F16G10.
A:Reference number: 221329
A:Accession: T13355
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-199 <GAT>
A:Cross-references: UNIPROT:O76588; UNIPARC:UPI000007891D; EMBL:AF077537; PIDN:AAC26272.
A:Experimental source: strain Bristol N2; clone F16G10
C:Genetics:
A:Gene: CESP:F16G10.4
A:Map position: 2
A:Introns: 56/3; 69/1; 89/1; 133/3
C:Superfamily: Caenorhabditis elegans hypothetical protein F16G10.4

Query Match 68.0%; Score 17; DB 2; Length 199;
Best Local Similarity 22.2%; Pred. No. 51;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 72 CEATSTATC 80

RESULT 7

S63342
hypothetical protein YNR014w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein N2060
C:Species: Saccharomyces cerevisiae
C>Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 09-Jul-2004
C:Accession: S63342; S63344; S63340
R:Mauter, C.T.C.; Urbanus, J.H.M.; Plant, R.U.
Submitted to the Protein Sequence Database, April 1996
A:Reference number: S63266

A:Accession: S63342
A:Molecule type: DNA
A:Residues: 1-212 <MAU>
A:Cross-references: UNIPROT:P53719; UNIPARC:UPI000013BAB3; EMBL:Z71629; NID:G1302493; PI
A:Experimental source: strain S288C
R:Ripstein, A.; Floeth, M.; Fritz, C.; Heuss-Neitzel, D.; Hilbert, H.; Moestl, D.
Submitted to the Protein Sequence Database, April 1996
A:Reference number: S62944
A:Accession: S63344
A:Molecule type: DNA
A:Residues: 1-212 <DUB>
A:Cross-references: UNIPARC:UPI000013BAB3; EMBL:Z71629; NID:G1302493; PID:G239786; PID:G

A:Experimental source: strain S288C
R:Aert, R.; Verhaesselt, P.; Voet, M.; Volckaert, G.
Submitted to the Protein Sequence Database, April 1996
A:Reference number: S62910
A:Accession: S63340
A:Molecule type: DNA
A:Residues: 1-26 <ER>
A:Cross-references: UNIPARC:UPI0000179C25; EMBL:Z71629; MIPS:YNR014w
A:Experimental source: strain S288C
C:Genetics:
A:Cross-references: SGD:S0005297
A:Map position: 14R

Query Match 68.0%; Score 17; DB 2; Length 212;
Best Local Similarity 22.2%; Pred. No. 52;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 40 CSSATSKAC 48

RESULT 8

C70812
probable lpgQ protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: C70812
R:Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; M01D:9829587; PMID:9634230
A:Accession: C70812
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-214 <COL>
A:Cross-references: UNIPROT:O53846; UNIPARC:UPI0000165256; GB:AL022004; GB:AL123456; NIT
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: lpgQ

Query Match 68.0%; Score 17; DB 2; Length 214;
Best Local Similarity 22.2%; Pred. No. 52;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 15 CAIAATTAC 23

RESULT 9

H72532
hypothetical protein APE2236 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: H72532
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai
awa, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;

DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
 A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: H72532
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-244 <RAM>
 A:Cross-references: UNIPROT:Q9Y9Q2; UNIPARC:UPI000005E27C; DDBJ:AP000063; NID:95105654;
 A:Experimental source: strain K1
 A:Genetics:
 A:Gene: APE2236
 C:Superfamily: Aeropyrum pernix hypothetical protein APE2236

Query Match 68.0%; Score 17; DB 2; Length 244;
 Best Local Similarity 22.2%; Pred. No. 54;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
 DB 107 CSSSSAALC 115

RESULT 10
 T40415
 hypothetical protein SPBC4.01 - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T40415
 R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Harris, D.
 submitted to the EMBL Data Library, October 1999
 A:Reference number: Z21927
 A:Accession: T40415
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-248 <MCD>
 A:Cross-references: UNIPROT:Q9US9; UNIPARC:UPI000006AF49; EMBL:AL121863; PIDN:CA858400;
 A:Experimental source: strain 972h-; cosmid c4
 C:Genetics:
 A:Gene: SPDB:SPBC4.01
 A:Map position: 2
 A:introns: 230/2

Query Match 68.0%; Score 17; DB 2; Length 248;
 Best Local Similarity 22.2%; Pred. No. 54;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
 DB 146 CLTTSTAAC 154

RESULT 11
 H71353
 conserved hypothetical protein TP0182 - syphilis spirochete
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
 C:Accession: H71353
 R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Ueberback, T.; Mcd
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 Science 281, 375-388, 1998
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
 A:Reference number: A71250; MUID:98332770; PMID:9665876
 A:Accession: H71353
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-259 <COL>
 A:Cross-references: UNIPROT:Q83212; UNIPARC:UPI000000CA51; GB:AE001202; GB:AE000520; NID
 A:Experimental source: strain Nichols
 C:Genetics:
 A:Gene: TP0182
 C:Superfamily: RNA-binding protein, YrDC type

Query Match 68.0%; Score 17; DB 2; Length 259;
 Best Local Similarity 22.2%; Pred. No. 55;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
 DB 31 CASASLSSC 39

RESULT 12
 AC0661
 probable lipoprotein STY1394 [imported] - Salmonella enterica subsp. enterica serovar Ty
 C:Species: Salmonella enterica subsp. enterica serovar Typh
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AC0661
 R:Fairhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 , S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AC0661
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-281 <PAR>
 A:Cross-references: UNIPARC:UPI0000059FOE; GB:AL513382; PIDN:CAD01660.1; PID:gl6502512;
 C:Genetics:
 A:Gene: STY1394

Query Match 68.0%; Score 17; DB 2; Length 281;
 Best Local Similarity 22.2%; Pred. No. 56;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
 DB 208 CAATAAEC 216

RESULT 13
 T48975
 xyloglucan endo-transglycosylase - Arabidopsis thaliana
 N:Alternate names: protein FlAD17.60
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
 C:Accession: T48975
 R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unsel, M.; Mewes, H.W.; Rudd, S.;
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z25008
 A:Accession: T48975
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-293 <JOR>
 A:Cross-references: UNIPROT:P93046; UNIPARC:UPI00000487CE; EMBL:AJ353992; GSPDB:GN00061.
 A:Experimental source: cultivar Columbia; BAC clone FlAD17
 C:Genetics:
 A:Gene: ATSP:FlAD17.60
 A:Map position: 3
 A:introns: 68/1; 101/3; 174/2
 C:Superfamily: endoxyloglucan transferase

Query Match 68.0%; Score 17; DB 2; Length 293;
 Best Local Similarity 22.2%; Pred. No. 57;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
 DB 238 CTADSSSC 246

RESULT 14
 T01948

hypothetical protein F1104.13 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004
C/Accession: T01948
R/Author: Thiele, J.; Stoneking, T.; Langston, Y.; Trevisan, E.
Submitted to the EMBL Data Library, October 1998
A/Description: The sequence of A. thaliana F1104.
A/Reference number: Z14466
A/Accession: T01948
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-298 <ABU>
A/Cross-references: UNIPROT:Q82599; UNIPARC:UP100000A5CDC; EMBL:AF096370; NID:93695372;
C/Genetics:
A/Map position: 4
A/Insertions: 158/3
A/Note: F1104.13
C/Superfamily: Arabidopsis thaliana hypothetical protein F21F14.120

Query Match 68.0%; Score 17; DB 2; Length 298;
Best Local Similarity 22.2%; Pred. No. 57;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9
DB 200 CXXXXXXXC 208

RESULT 15
A85019
probable bHLH DNA-binding protein [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C/Accession: A85019
R/Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A/Reference number: A85001; MUID:20083488; PMID:10617198
A/Accession: A85019
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-315 <STO>
A/Cross-references: UNIPROT:Q9M128; UNIPARC:UP1000009DFA; GB:NC_001268; NID:97268189; F
C/Genetics:
A/Map position: 4
A/Insertions: 174/01460
C/Superfamily: Arabidopsis thaliana hypothetical protein F21F14.120

Query Match 68.0%; Score 17; DB 2; Length 315;
Best Local Similarity 22.2%; Pred. No. 58;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9
DB 188 CXXXXXXXC 196

RESULT 16
AB1224
cobalamin biosynthesis protein Cbid homolog cbid [imported] - Listeria monocytogenes (str
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AB1224
R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fehli, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunz, F.; Kurapk, G.; Madueno, E.; Maltournam, A.; Ma
Ok, C.; Schleuter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AB1224

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-373 <GLA>
A/Cross-references: UNIPROT:Q8Y757; UNIPARC:UP10000054CCA; GB:NC_003210; PIDN:CAC9272.1
C/Genetics:
A/Experimental source: strain EGD-e
C/Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 68.0%; Score 17; DB 2; Length 373;
Best Local Similarity 22.2%; Pred. No. 61;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9
DB 21 CAAAAKAC 29

RESULT 17
AD1577
cobalamin biosynthesis protein Cbid homolog cbid [imported] - Listeria innocua (strain
C/Species: Listeria innocua
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AD1577
R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fehli, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunz, F.; Kurapk, G.; Madueno, E.; Maltournam, A.; Ma
Ok, C.; Schleuter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AD1577
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-373 <GLA>
A/Cross-references: UNIPROT:Q92C14; UNIPARC:UP100000CC4A4; GB:AL592022; PIDN:CAC96388.1
C/Genetics:
A/Experimental source: strain C1p11262
C/Genetics:
A/Map position: 4
C/Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 68.0%; Score 17; DB 2; Length 373;
Best Local Similarity 22.2%; Pred. No. 61;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9
DB 21 CAAAAKAC 29

RESULT 18
T43352
nuclear receptor NHR-13 - Caenorhabditis elegans (fragment)
C/Species: Caenorhabditis elegans
C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C/Accession: T43352
R/Sluder, A.E.; Mathews, S.W.; Hough, D.; Yin, V.P.; Maina, C.V.
Genome Res. 9, 103-120, 1999
A/Title: The nuclear receptor superfamily has undergone extensive proliferation and dive
A/Reference number: Z22443; MUID:99148134; PMID:10022975
A/Accession: T43352
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-415 <SLU>
A/Cross-references: UNIPROT:Q9XYB8; UNIPARC:UP100000755B6; EMBL:AF083228; NID:94139081;
C/Genetics:
A/Map position: 5
A/Insertions: nhr-13
A/Note: nhr-13
Query Match 68.0%; Score 17; DB 2; Length 415;
Best Local Similarity 22.2%; Pred. No. 63;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
Db 12 CSSSSNSC 20

RESULT 19

T33024
hypothetical protein K07H8.2 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C/Accession: T33024
R/Fulton, B.; Hawkins, J.; Gattung, S.; Wohldmann, P.; Elliott, G.

A/Description: The sequence of C. elegans cosmid K07H8.

A/Reference number: Z21264

A/Accession: T33024

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-507 <FUL>
A/Cross-references: UNIPROT:O45182; UNIPARC:UPI0000080E7E; EMBL:AF047659; PIDN:AAC04427.

A/Experimental source: strain Bristol N2; clone K07H8

C/Genetics:

A/Genes: CESP:K07H8.2

A/Map position: 4

A/Introns: 20/3; 55/2; 93/3; 152/3; 447/3

Query Match 68.0%; Score 17; DB 2; Length 507;
Best Local Similarity 22.2%; Pred. No. 67;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
Db 210 CASSLATAC 218

RESULT 20

T48004
multifunctional aminoacyl-tRNA ligase-like protein - Arabidopsis thaliana

N/Alternate names: protein T17J13.80

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004

C/Accession: T48004

R/Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Lemcke, K.; Mayer, submitted to the Protein Sequence Database, February 2000

A/Reference number: Z24482

A/Accession: T48004

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-530 <RIE>

A/Cross-references: UNIPROT:Q9M1R2; UNIPARC:UPI000009CD43; EMBL:AL138651

A/Experimental source: cultivar Columbia; BAC clone T17J13

C/Genetics:

A/Map position: 3

A/Introns: 64/3; 94/3; 133/3; 196/2; 218/3; 259/3; 310/3; 405/3; 478/3; 510/1

A/Note: T17J13.80

C/Superfamily: proline-tRNA ligase pros

Query Match 68.0%; Score 17; DB 2; Length 530;
Best Local Similarity 22.2%; Pred. No. 67;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
Db 354 CTATASALC 362

RESULT 21

A42143
Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 09-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C/Accession: A42143; T29535

R/Bowerman, B.; Eaton, B.A.; Priess, J.R.

Cell 68, 1061-1075, 1992

A/Title: skn-1, a maternally expressed gene required to specify the fate of ventral bias

A/Reference number: A42143; MUID:92191285; PMID:1547503

A/Accession: A42143

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-533 <BOW>

A/Cross-references: UNIPROT:O8MPW3; UNIPARC:UPI0000081925

A/Note: sequence extracted from NCBI backbone (NCBIN:88973, NCBI:88974)

R/Pauley, A.; Gattung, S.

submitted to the EMBL Data Library, December 1995

A/Description: The sequence of C. elegans cosmid T19E7.

A/Reference number: Z20637

A/Accession: T29535

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-386 <PAU>

A/Cross-references: UNIPARC:UPI000017A521; EMBL:U42843; PIDN:AAA83594.1

C/Genetics:

A/Genes: CESP:skn-1

A/Introns: 141/3; 249/1; 300/1

C/Keywords: nucleus

Query Match 68.0%; Score 17; DB 2; Length 533;
Best Local Similarity 22.2%; Pred. No. 67;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
Db 384 CTTSSSSTC 392

RESULT 22

T07739

probable ferredoxinase (EC 4.99.1.1) - potato

C/Species: Solanum tuberosum (potato)

C/Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004

C/Accession: T07739

R/Johnston, D.J.

submitted to the EMBL Data Library, April 1998

A/Reference number: Z15932

A/Accession: T07739

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-543 <JOH>

A/Cross-references: UNIPROT:O64391; UNIPARC:UPI00000A543B; EMBL:AJ005802; PIDN:CAA06705

A/Experimental source: cv. Bintje

C/Function:

A/Description: catalyzes the insertion of iron into protoporphyrin to produce heme

A/Pathway: heme biosynthesis

C/Superfamily: ferredoxinase

C/Keywords: iron; lyase; mitochondrial inner membrane

Query Match 68.0%; Score 17; DB 2; Length 543;
Best Local Similarity 22.2%; Pred. No. 67;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
Db 31 CTTSSPASAC 39

RESULT 23

THH0B

thrombomodulin precursor [validated] - human

C/Species: Homo sapiens (man)

C/Date: 31-Dec-1988 #sequence_revision 12-May-1995 #text_change 09-Jul-2004

C/Accession: A41442; A28307; A29680; A27073; JX0264; S38954

R/Shirai, T.; Shiojiri, S.; Ito, H.; Yamamoto, S.; Kusumoto, H.; Deyashiki, Y.; Maruyama

U. Biochem. 103, 281-285, 1988

A/Title: Gene structure of human thrombomodulin, a cofactor for thrombin-catalyzed activ

A:Reference number: A41442; MWID:88227901; PMID:2836377
A:Accession: A41442
A:Molecule type: DNA
A:Residues: 1-575 <SUZ>
A:Cross-references: UNIPROT:P07204; UNIPARC:UPI00000498EB; DDBJ:DD0210; NID:9220126; PID:R:Jackman, R.W.; Beiler, D.L.; Fritze, L.; Soff, G.; Rosenberg, R.D.
Proc. Natl. Acad. Sci. U.S.A. 84, 6425-6429, 1987
A:Title: Human thrombomodulin gene is intron depleted: nucleic acid sequences of the cDN
A:Reference number: A28307; MWID:87317665; PMID:2819876
A:Accession: A28307
A:Molecule type: DNA; mRNA
A:Residues: 1-472; 'A', 474-575 <JAC>
A:Cross-references: UNIPARC:UPI00000002BD; GB:J02973; NID:9339658; PIDN:AAA61175.1; PID:R:Suzuki, K.; Kusumoto, H.; Deysashiki, Y.; Nishiooka, J.; Maruyama, I.; Zushi, M.; Kawaba
EMBO J. 6, 1891-1897, 1987
A:Title: Structure and expression of human thrombomodulin, a thrombin receptor on endoth
A:Reference number: A29680; MWID:88004395; PMID:2820710
A:Accession: A29680
A:Molecule type: mRNA
A:Residues: 1-575 <SUZ>
A:Cross-references: UNIPARC:UPI00000498PB; GB:X05495; NID:937123; PIDN:CAA29045.1; PID:G
A:Experimental source: lung endothelium
A:Note: part of this sequence, including the amino end of the mature protein, were deter
R.Wen, D.; Dittman, W.A.; Ye, R.D.; Deaven, L.L.; Majerus, P.W.; Sadler, J.E.
Biochemistry 26, 4350-4357, 1987
A:Title: Human thrombomodulin: complete cDNA sequence and chromosome localization of the
A:Reference number: A27073; MWID:88024950; PMID:2822087
A:Accession: A27073
A:Molecule type: mRNA
A:Residues: 1-472; 'A', 474-575 <WEN>
A:Cross-references: UNIPARC:UPI00000002BD; GB:M16552; NID:9339656; PIDN:AAB59508.1; PID:
A:Experimental source: placenta
A:Note: parts of this sequence were determined by protein sequencing
R.Yamamoto, S.; Mizoguchi, T.; Tamaki, T.; Ohkuchi, M.; Kimura, S.; Aoki, N.
J. Biochem. 113, 433-440, 1993
A:Title: Urinary thrombomodulin, its isolation and characterization.
A:Reference number: JX0264; MWID:93293792; PMID:8390446
A:Accession: JX0264
A:Molecule type: protein; mRNA
A:Residues: 19-472; 'A', 474-486 <YAM>
A:Cross-references: UNIPARC:UPI0000173341
A:Experimental source: urine
A:Note: the urinary form appears to be identical with that circulating in plasma
R.Gerlitz, B.; Hassell, T.; Vlahos, C.J.; Parkinson, J.F.; Bang, N.U.; Grinnell, B.W.
Biochem. J. 295, 131-140, 1993
A:Title: Identification of the predominant glycosaminoglycan-attachment site in soluble
serine.
A:Reference number: S38954; MWID:94029900; PMID:8216207
A:Accession: S38954
A:Molecule type: protein
A:Residues: 475-491; 'X', 493-494 <GER>
A:Cross-references: UNIPARC:UPI0000173342
A:Note: the residue designated 'X' was determined to be a Ser with covalently bound chon
R.Meilinger, D.P.; Komives, E.A.
submitted to the Brookhaven Protein Data Bank, September 1995
A:Reference number: A67369; PDB:1ZAO
A:Contents: annotation, conformation and disulfide bond assignments by (1)H-NMR, residue
R.Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, August 1994
A:Reference number: A52804; PDB:1HLT
A:Contents: X-ray crystallography, 3.0 angstroms, residues 426-442
R.Hirabai, R.; Komives, E.A.; Ni, F.
submitted to the Brookhaven Protein Data Bank, November 1995
A:Reference number: A65583; PDB:1FGD
A:Contents: annotation, conformation by (1)H-NMR, residues 427-444
Protein Sci. 5, 195-203, 1996
A:Title: Structural resiliency of an EGF-like subdomain bound to its target protein, the
A:Reference number: A58595; MWID:96276211; PMID:8745536
A:Contents: annotation, conformation by (1)H-NMR
C:Genetics:
A:Gene: GDB:THBD
A:Cross-references: GDB:119613; OMIM:188040

A:Map position: 20p11.2-20p11.2
A:introns: #status absent
C:Complex: homodimer, urinary form
C:Function:
A:Description: inhibits thrombin activation of fibrinogen; cofactor for thrombin activat
A:Pathway: blood coagulation/moderation
A:Note: the membrane-bound form is located on the endothelium/luminal surface of arterie
C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology
C:Keyword: anticoagulant; beta-hydroxyaspartate; beta-hydroxyaspartic acid; blood coag
e protein
F:1-16/Domain: signal sequence #status predicted <SIG>
F:19-575/Product: thrombomodulin, membrane-bound form #status predicted <MAT>
F:19-513/Domain: extracellular #status predicted <EXT>
F:124-167/Domain: C-type lectin homology, urinary form #status experimental <MAU>
F:177-199/Region: PEST sequence
F:201-233/Region: PEST sequence
F:245-280/Domain: EGF homology <EG1>
F:288-323/Domain: EGF homology <EG2>
F:329-362/Domain: EGF homology <EG3>
F:369-404/Domain: EGF homology <EG4>
F:408-439/Domain: EGF homology <EG5>
F:445-480/Domain: EGF homology <EG6>
F:485-513/Region: PEST sequence
F:517-539/Domain: intracellular #status predicted <INT>
F:540-575/Domain: transmembrane #status predicted <TMN>
F:47,115,116,382,409/Binding site: carboxylate (Asn) (covalent) #status predicted
F:174,225,411,504/Binding site: carboxylate (Thr) (covalent) #status predicted
F:245-256,252-265,267-280,286-296,292-308,310-323,329-340,336-349,351-362,369-378,374-38
F:334,498/Binding site: carboxylate (Ser) (covalent) #status predicted
F:342/Modified site: erythro-beta-hydroxyaspartate (Asn) #status experimental
F:90,492/Binding site: chondroitin sulfate (Ser) (covalent) (partial) #status experimen

Query Match 68.0%; Score 17; DB 1; Length 575;
Best Local Similarity 22.2%; Pred. No. 69;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 280 CTRASTQC 288

RESULT 24
T07953
Lectin-like protein zsp2, zygote-specific - Chlamydomonas reinhardtii
C:Species: Chlamydomonas reinhardtii
C:date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004
C:Accession: T07953
R:Suzuki, L.; Woessner, J.P.; Uchida, H.; Kuriyama, H.; Takano, H.; Kuroiwa, H.; Yuasa,
submitted to the EMBL Data Library, March 1998
A:Description: A lectin-like protein mediates the assembly of the extracellular matrix c
A:Reference number: Z16243
A:Accession: T07953
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-582 <SUZ>
A:Cross-references: UNIPROT:065153; UNIPARC:UPI00000454CF; EMBL:AF053099; NID:92997677;
C:Genetics:
A:Gene: zsp2
A:introns: 27/1, 292/3

Query Match 68.0%; Score 17; DB 2; Length 582;
Best Local Similarity 22.2%; Pred. No. 69;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 257 CARSSTTC 265

RESULT 25
T07952

lectin-like protein, zygote-specific - Chlamydomonas reinhardtii
C:Species: Chlamydomonas reinhardtii
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004
C:Accession: T07952
R:Suzuki, L.; Woessner, J.P.; Uchida, H.; Kuriyama, H.; Takano, H.; Kuroiwa, H.; Yuasa, R.
submitted to the EMBL Data Library, March 1998
A:Description: A lectin-like protein mediates the assembly of the extracellular matrix
A:Reference number: Z16243
A:Accession: T07952
A:Status: preliminary; translated from GB/EMBL/DDBL
A:Molecule type: DNA
A:Residues: 1-582 <SU2>
A:Cross-references: UNIPROT:Q9SB11; UNIPARC:UPI00000454C; EMBL:AF053098; NID:g2997675;
C:Genetics:
A:Gene: zsp2
A:Introns: 27/1; 292/3

Query Match 68.0%; Score 17; DB 2; Length 582;
Best Local Similarity 22.2%; Pred. No. 69;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
DB 257 CARSTTTC 265

RESULT 26
B25682
homeotic protein Engrailed - fruit fly (Drosophila virilis)
C:Species: Drosophila virilis
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004
C:Accession: B25682
R:Kasatis, J.A.; Poole, S.J.; Wright, D.K.; O'Farrell, P.H.
EMBO J. 5, 3583-3589, 1986
A:title: Sequence conservation in the protein coding and intron regions of the engrailed
A:Reference number: A91059; MUID:87161768; PMID:2881781
A:Accession: B25682
A:Molecule type: DNA
A:Residues: 1-584 <KAS>
A:Cross-references: UNIPROT:P09145; UNIPARC:UPI000012CA14; GB:X04727; NID:g9173; PIDN:CA
C:Genetics:
A:Gene: en
A:Cross-references: FlyBase:FBgn0013111
A:Introns: 470/1; 502/3
C:Superfamily: engrailed homeotic protein; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:487-543/Domain: homeobox homology <Hox>

Query Match 68.0%; Score 17; DB 2; Length 584;
Best Local Similarity 22.2%; Pred. No. 69;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
DB 384 CSASSSSGC 392

RESULT 27
S41011
hypothetical protein ZK757.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S41011
R:Thomas, K.
submitted to the EMBL Data Library, December 1993
A:Reference number: S41011
A:Accession: S41011
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-589 <THO>
A:Cross-references: UNIPROT:P34679; UNIPARC:UPI000013CC25; EMBL:Z29121; NID:g438366; PID
C:Genetics:
A:Introns: 19/2; 54/3; 114/3; 146/3; 178/3; 209/1; 253/1; 302/3; 324/3; 352/1; 392/3; 46

C:Keywords: transmembrane protein
Query Match 68.0%; Score 17; DB 2; Length 589;
Best Local Similarity 22.2%; Pred. No. 69;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
DB 100 CITSSSTSC 108

RESULT 28
A42086
CD30 antigen precursor - human
N:Alternate names: Ki-1 antigen; nerve growth factor receptor family member CD30
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: A42086
R:Durkop, H.; Latza, U.; Hummel, M.; Eitelbach, F.; Seed, B.; Stein, H.
Cell 68, 421-427, 1992
A:title: Molecular cloning and expression of a new member of the nerve growth factor rec
A:Reference number: A42086; MUID:92154659; PMID:1310894
A:Accession: A42086
A:Molecule type: mRNA
A:Residues: 1-595 <DUR>
A:Cross-references: UNIPROT:P28908; UNIPARC:UPI000000971; GB:M83554; NID:g180095; PIDN:
A:Experimental source: HUT-102 cell line
A>Note: sequence extracted from NCBI backbone (NCBIN:82088, NCBI:P:82090)
C:Genetics:
A:Gene: GDB:CD30; DIS166E
A:Cross-references: GDB:11547; OMIM:153243
A:Map position: 1p36-1p36
C:Keywords: glycoprotein; growth factor receptor; transmembrane protein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-383/Domain: extracellular #status predicted <EXT>
F:384-407/Domain: transmembrane #status predicted <TM>
F:408-595/Domain: intracellular #status predicted <CTT>
F:101,276/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 68.0%; Score 17; DB 2; Length 595;
Best Local Similarity 22.2%; Pred. No. 69;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
DB 289 CATSATNSC 297

RESULT 29
D83286
hypothetical protein PA2886 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: D83286
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lapidis, K.; Lam,
-; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83286
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-600 <STO>
A:Cross-references: UNIPROT:Q9H2M1; UNIPARC:UPI00000C5861; GB:AE004714; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2886

Query Match 68.0%; Score 17; DB 2; Length 600;
Best Local Similarity 22.2%; Pred. No. 69;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 92 CAALAAAC 100

RESULT 30

T15408
hypothetical protein C04F6.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #ext_change 09-Jul-2004
C:Accession: T15408
R:Nhan, M.
submitted to the EMBL Data Library, December 1995
A:Description: The sequence of C. elegans cosmid C04F6.
A:Reference number: Z18346

A:Accession: T15408
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-617 <NHA>
A:Cross-references: UNIPROT:Q11174; UNIPARC:UPI000060EB1; EMBL:U42835; NID:g1125760; P
C:Genetics:
A:Gene: CRSP:C04F6.3
A:introns: 28/1; 66/2; 504/1

Query Match 68.0%; Score 17; DB 2; Length 617;
Best Local Similarity 22.2%; Pred. No. 70;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 524 CTTTASGC 532

RESULT 31

A36962
laccase (EC 1.10.3.2) precursor - fungus (*Filobasidium floriforme*) (ATCC 34873)
N:Alternate names: diphenol oxidase
C:Species: *Filobasidium neoformanis*, *Cryptococcus neoformans*
C>Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #ext_change 31-Dec-2004
C:Accession: A36962
R:Williamson, P.R.
J. Bacteriol. 176, 656-664, 1994

A:Title: Biochemical and molecular characterization of the diphenol oxidase of *Cryptococ*
A:Reference number: A36962; MUID:94131944; PMID:8300520
A:Accession: A36962
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-624 <WIL>
A:Cross-references: UNIPROT:Q9UQZ7; UNIPARC:UPI00001751D2; GB:I22866
C:Genetics:
A:Gene: CNLACT1
A:introns: 22/3; 69/3; 117/1; 137/3; 176/2; 204/2; 223/3; 238/3; 371/2; 426/3; 55
C:Keywords: copper; glycoprotein; oxidoreductase

Query Match 68.0%; Score 17; DB 2; Length 624;
Best Local Similarity 22.2%; Pred. No. 70;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 243 CTATGSSSC 251

RESULT 32

A41029
integrin beta-8 chain precursor - human
C:Species: *Homo sapiens* (man)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #ext_change 31-Dec-2004
C:Accession: A41029
R:Moyle, M.; Napier, M.A.; McLean, J.W.
J. Biol. Chem. 266, 19650-19658, 1991
A:Title: Cloning and expression of a divergent integrin subunit beta-8.
A:Reference number: A41029; MUID:92011767; PMID:1918072

A:Accession: A41029
A:Molecule type: mRNA
A:Residues: 1-769 <MOY>
A:Cross-references: UNIPROT:P26012; UNIPARC:UPI000012DA14; GB:M73780; NID:g184520; PID:
C:Superfamily: integrin, beta subunit; laminin-type EGF-like homology
C:Keywords: cell adhesion; cytoskeleton; duplication; extracellular matrix; glycoprotein.

Query Match 68.0%; Score 17; DB 2; Length 769;
Best Local Similarity 22.2%; Pred. No. 74;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 47 CASSNAAC 55

RESULT 33

A46140
diacylglycerol kinase (EC 2.7.1.107) - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C>Date: 22-Nov-1993 #sequence_revision 18-Nov-1994 #ext_change 09-Jul-2004
C:Accession: A46140
R:Masai, I.; Hosoya, T.; Kojima, S.; Hotta, Y.
Proc. Natl. Acad. Sci. U.S.A. 89, 6030-6034, 1992
A:Title: Molecular cloning of a *Drosophila* diacylglycerol kinase gene that is expressed
A:Reference number: A46140; MUID:92335231; PMID:1321433
A:Accession: A46140
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-791 <MAS>
A:Cross-references: UNIPROT:Q01583; UNIPARC:UPI000016BB61; GB:D11120; NID:g217332; PID:
A:Note: sequence extracted from NCBI backbone (NCBIN:108569; NCBI:P:108571)
C:Genetics:

A:Gene: FlyBase:Dgk
A:Cross-references: FlyBase:FBgn0004568
C:Superfamily: fruit fly diacylglycerol kinase; calmodulin repeat homology; protein kin
C:Keywords: EF hand; phosphotransferase

Query Match 68.0%; Score 17; DB 2; Length 791;
Best Local Similarity 22.2%; Pred. No. 75;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 18 CATAAAVAC 26

RESULT 34

B46140
diacylglycerol kinase (EC 2.7.1.107) - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #ext_change 29-Sep-1999
C:Accession: B46140
R:Masai, I.; Hosoya, T.; Kojima, S.; Hotta, Y.
Proc. Natl. Acad. Sci. U.S.A. 89, 6030-6034, 1992
A:Title: Molecular cloning of a *Drosophila* diacylglycerol kinase gene that is expressed
A:Reference number: A46140; MUID:92335231; PMID:1321433
A:Accession: B46140
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-796 <MAS>
A:Cross-references: UNIPARC:UPI00001756B6; GB:D11120
C:Genetics:

A:Gene: FlyBase:Dgk
A:Cross-references: FlyBase:FBgn0004568
C:Superfamily: fruit fly diacylglycerol kinase; calmodulin repeat homology; protein kina
C:Keywords: EF hand; phosphotransferase

Query Match 68.0%; Score 17; DB 2; Length 796;
Best Local Similarity 22.2%; Pred. No. 75;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

Db 18 CATAAAVAC 26

RESULT 35

A38308

Integrin beta-5 chain precursor - human

C:Species: Homo sapiens (man)

C>Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 31-Dec-2004

C:Accession: A38308; A35775; S12534; S11708

R:McClean, J.W.; Vestal, D.J.; Chersesh, D.A.; Bodary, S.C.

J. Biol. Chem. 265, 17126-17131, 1990

A>Title: cDNA sequence of the human integrin beta-5 subunit.

A:Reference number: A38308; UID:91009141; PMID:2211615

A:Accession: A38308

A:Molecule type: mRNA

A:Residues: 1-799 <MCL>

A:Cross-references: UNIPROT:P18084; UNIPARC:UPI00004D05B; GB:J05633; NID:g186504; PIDN:

A>Note: parts of this sequence, including the amino end of the mature protein, were conf

R:Suzuki, S.; Huang, Z.S.; Tanihara, H.

Proc. Natl. Acad. Sci. U.S.A. 87, 5354-5358, 1990

A>Title: Cloning of an integrin beta subunit exhibiting high homology with integrin beta

A:Reference number: A35775; UID:90319111; PMID:2371275

A:Accession: A35775

A:Molecule type: mRNA

A:Residues: 1-192 'A', 194-644, 'L', 646-789, 793-799 <SUZ>

A:Cross-references: UNIPARC:UPI000016ADB; GB:M35011; NID:g184524; PIDN:AAA5707.1; PID:

R:Ramakswamy, H.; Hemler, M.E.

EMBO J. 9, 1561-1568, 1990

A>Title: Cloning, primary structure and properties of a novel human integrin beta subunit

A:Reference number: S12534; UID:90228356; PMID:2328726

A:Accession: S12534

A:Molecule type: mRNA

A:Residues: 1-644, 'L', 646-799 <RAM>

A:Cross-references: UNIPARC:UPI000012DA10; EMBL:X53002; NID:g33952; PIDN:CAA37188.1; PID:

C:Genetics:

A:Gene: GDB:ITGB5

A:Cross-references: GDB:128005; OMIM:147561

A:Map position: 17q11-17qter

C:Superfamily: Integrin, beta subunit; laminin-type EGF-like homology

C:Keywords: cell adhesion; cytoskeleton; duplication; extracellular matrix; glycoprotein

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-799/Product: integrin beta-5 chain #status experimental <MAT>

F:25-719/Domain: extracellular #status predicted <EXT>

F:463-513/Domain: laminin-type EGF-like homology <LEG>

F:720-742/Domain: transmembrane #status predicted <TM>

F:743-799/Domain: intracellular #status predicted <INT>

F:347,460,477,505,552,586,654,705/Binding site: carbohydrate (asn) (covalent) #status pr

Query Match 68.0%; Score 17; DB 2; Length 799;

Best Local Similarity 22.2%; Pred. No. 75;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9

Db 28 CTSGSATSC 36

RESULT 36

T05201

hypothetical protein F410.140 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C:Accession: T05201

R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirksen, W.; Stiekema, W.; Hohnel, J.; New

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15402

A:Accession: T05201

A:Molecule type: DNA

A:Residues: 1-934 <BEV>

A:Cross-references: UNIPROT:Q5SMY8; UNIPARC:UPI0000098D1; EMBL:AL035525

A:Experimental source: cultivar Columbia; BAC clone F410

C:Genetics:

A:Map position: 4
A:introns: 123/2; 203/3; 248/1; 281/2; 347/2; 437/3; 470/3; 500/2; 533/3; 590/3; 641/3;
A>Note: F410.140

Query Match 68.0%; Score 17; DB 2; Length 934;

Best Local Similarity 22.2%; Pred. No. 79;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9

Db 701 CUSATATSC 709

RESULT 37

T41932

hypothetical protein U30 - human herpesvirus 7 (strain J1)

C:Species: human herpesvirus 7

A:Variety: strain J1

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T41932

R:Nicholas, J.

submitted to the EMBL Data Library, December 1995

A:Description: Determination and analysis of the complete nucleotide sequence of human h

A:Reference number: Z22022

A:Accession: T41932

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-938 <NIC>

A:Cross-references: UNIPROT:P52438; UNIPARC:UPI00000007D8; EMBL:U43400; PIDN:AA054692.1

A:Experimental source: strain J1

C:Genetics:

A>Note: U30

Query Match 68.0%; Score 17; DB 2; Length 938;

Best Local Similarity 22.2%; Pred. No. 79;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9

Db 689 CUSTLATTC 697

RESULT 38

T15881

hypothetical protein D1044.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15881

R:Pauley, A.

submitted to the EMBL Data Library, June 1994

A:Description: The sequence of C. elegans comid D1044.

A:Reference number: Z18423

A:Accession: T15881

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1895 <PAU>

A:Cross-references: UNIPARC:UPI000004CAE8; EMBL:U00065; NID:g495681; PID:g495684; PIDN:A

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:D1044.3

A:introns: 35/2; 63/2; 150/1; 191/3; 207/2; 405/3; 551/3; 629/1; 764/1; 933/1; 959/2; 10

Query Match 68.0%; Score 17; DB 2; Length 1895;

Best Local Similarity 22.2%; Pred. No. 96;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9

Db 1233 CTSSTSSSC 1241

RESULT 39

S50820

RESULT 44
H66753
prophage p12 protein 25 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: H66753
R:Boletín, A.; Wincker, P.; Mueger, S.; Jallón, O.; Malarre, K.; Weisenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s8
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: H66753
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-87 <STO>
A:Cross-references: UNIPROT:Q9CGR4; UNIPARC:UPI00000C6978; GB:AE005176; PID:g12723983; F
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: p1225

Query Match 64.0%; Score 16; DB 2; Length 87;
Best Local Similarity 22.2%; Pred. No. 1.1e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 30 CTAKSSLSC 38

RESULT 45
JS0036
Clara cell 10K protein precursor - human
N:Alternate names: urinary protein 1
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: JS0036; PS0309; A56890; I38397
R:Singh, G.; Katyal, S.L.; Brown, W.E.; Phillips, S.; Kennedy, A.L.; Anthony, J.; Squeglia
Biochim. Biophys. Acta 950, 329-337, 1998
A:Title: Amino-acid and cDNA nucleotide sequences of human Clara cell 10KDa protein.
A:Reference number: JS0036; MUID:89000784; PMID:3167058
A:Accession: JS0036
A:Molecule type: mRNA
A:Residues: 1-91 <SIN>
A:Cross-references: UNIPROT:P11684; UNIPARC:UPI00000369D1; GB:X13197; NID:g21131; PIDN:C
A:Accession: PS0309
A:Molecule type: protein
A:Residues: 22-23, 'X', 25-28, 'X', 30-31, 'X', 33-36 <S2>
A:Cross-references: UNIPARC:UPI0000177C39
R:Bernard, A.; Reels, H.; Lauwery, R.; Witters, R.; Gielens, C.; Soumilion, A.; Van De
Clin. Chim. Acta 207, 239-249, 1992
A:Title: Human urinary protein 1: evidence for identity with the Clara cell protein and
A:Reference number: A56890; MUID:93009001; PMID:1395029
A:Accession: A56890
A:Molecule type: protein
A:Residues: 22-45 <BER>
A:Cross-references: UNIPARC:UPI0000036975
A:Experimental source: urine
A:Note: Sequence extracted from NCBI backbone (NCBIP:119391)
R:Hay, J.G.; Daniel, C.; Chu, C.; Crystal, R.G.
Am. J. Physiol. 268, 565-575, 1995
A:Title: Human CC10 gene expression in airway epithelium and subchromosomal locus sugges
A:Reference number: I38397
A:Accession: I38397
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-91 <RES>
A:Cross-references: UNIPARC:UPI00000369D1; EMBL:U01101; NID:g457932; PIDN:AAA81885.1; PI
C:Comment: This protein consists of two identical polypeptides linked by two disulfide b
C:Genetics:
A:Gene: CC10
C:Superfamily: uteroglobin
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-91/Product: Clara cell 10K protein #status experimental <MAT>

Query Match 64.0%; Score 16; DB 2; Length 91;
Best Local Similarity 22.2%; Pred. No. 1.1e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 16 CSSASSDIC 24

RESULT 46
UGMS
uteroglobin precursor - mouse
N:Alternate names: CC10; Clara cell 10K protein precursor; Clara cell secretory protein
C:Species: Mus musculus (house mouse)
C:Date: 03-May-1994 #sequence_revision 21-Jan-1997 #text_change 09-Jul-2004
C:Accession: A53025; A56656; I51925; S24783
R:Stripp, B.R.; Hultman, J.A.; Bohinski, R.J.
Genomics 20, 27-35, 1994
A:Title: Structure and regulation of the murine Clara cell secretory protein gene.
A:Reference number: A53025; MUID:94292183; PMID:8020953
A:Accession: A53025
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-96 <STR>
A:Cross-references: UNIPROT:Q06318; UNIPARC:UPI00000011D0; GB:L24372; NID:g461147; PIDN:
R:Singh, G.; Katyal, S.L.; Brown, W.E.; Kennedy, A.L.
Exp. Lung Res. 19, 67-75, 1993
A:Title: Mouse Clara cell 10-KDa (CC10) protein: cDNA nucleotide sequence and molecular
A:Reference number: A56656; MUID:93178380; PMID:8440203
A:Accession: A56656
A:Molecule type: mRNA; protein
A:Residues: 1-96 <SIN>
A:Cross-references: UNIPARC:UPI00000011D0; EMBL:X67702; NID:g49690; PIDN:CAA47936.1; PID
A:Experimental source: lung
A:Note: Sequence extracted from NCBI backbone (NCBIP:126148)
A:Note: parts of this sequence, including the amino end of the mature protein, were conf
R:Marraf, L.R.; Finegold, M.J.; Stanley, L.A.; Major, A.; Watkins, H.K.; DeMayo, F.J.
Am. J. Respir. Cell Mol. Biol. 9, 231-238, 1993
A:Title: Cloning and tissue-specific expression of the cDNA for the mouse Clara cell 10K
A:Reference number: I51925; MUID:94000840; PMID:8398159
A:Accession: I51925
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-96 <RES>
A:Cross-references: UNIPARC:UPI00000011D0; GB:L04503; NID:g20213; PIDN:AAA03625.1; PID
C:Genetics:
A:Intons: 19/1; 81/3
C:Complex: homodimer linked by two disulfide bonds
C:Superfamily: uteroglobin
C:Keywords: lung; steroid binding; uterus
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-96/Product: uteroglobin #status predicted <MAT>
F:24/Disulfide bonds: interchain (to 90) #status predicted
F:90/Disulfide bonds: interchain (to 24) #status predicted

Query Match 64.0%; Score 16; DB 1; Length 96;
Best Local Similarity 22.2%; Pred. No. 1.2e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 16 CSSASSDIC 24

RESULT 47
A36581
polychlorinated biphenyl-binding protein precursor - rat
N:Alternate names: Clara cell 10K secretory protein
C:Species: Rattus norvegicus (Norway rat)
C:Date: 08-Mar-1991 #sequence_revision 08-Mar-1991 #text_change 09-Jul-2004
C:Accession: A36581; S10185; S21676
R:Nordlund-Moeller, L.; Andersson, O.; Ahlgren, R.; Schilling, J.; Gällner, M.; Gustafss

J. Biol. Chem. 265, 12690-12693, 1990
 A:Title: Cloning, structure, and expression of a rat binding protein for polychlorinated
 A:Reference number: A36581, MUID:9034266; PMID:2115524
 A:Accession: A36581
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-96 <NOR>
 A:Cross-references: UNIPROT:P17559; UNIPARC:UPI0000112E8B; GB:J05536; NID:g206039; PIDN:
 R:Ragen, G.; Wolf, M.; Karyal, S.L.; Singh, G.; Beato, M.; Suske, G.
 Nucleic Acids Res. 18, 2939-2946, 1990
 A:Title: Tissue-specific expression, hormonal regulation and 5'-flanking gene region of
 A:Reference number: S10185, MUID:90272398; PMID:2349092
 A:Accession: S10185
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-18 <HNG>
 A:Cross-references: UNIPARC:UPI00001707CC; EMBL:X51318; NID:g55536; PIDN:CA35701.1; PID
 R:Umland, T.C.; Swaminathan, S.; Furey, W.; Singh, G.; Pletcher, J.; Sax, M.
 J. Mol. Biol. 224, 441-448, 1992
 A:Title: Refined structure of rat Clara cell 17 kDa protein at 3.0 Å resolution.
 A:Reference number: S21676; MUID:92219263; PMID:1560460
 A:Content: annotation, X-ray crystallography, 3.0 angstroms
 C:Superfamily: uteroglobin
 F:24/Disulfide bonds: interchain (to 24) #status experimental
 F:90/Disulfide bonds: interchain (to 24) #status experimental

Query Match 64.0%; Score 16; DB 2; Length 96;
 Best Local Similarity 22.2%; Pred. No. 1.2e+02;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
 DB 16 CSSASSDIC 24

RESULT 48
 A10334
 conserved hypothetical protein YPO2745 [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
 C:Accession: A10334
 R:Parhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
 deo-Taragga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: A10334
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-98 <KUR>
 A:Cross-references: UNIPROT:Q8ZD47; UNIPARC:UPI00000CD967; GB:AL590842; PIDN:CAC92984.1;
 C:genetics:
 A:Gene: YPO2745
 C:Superfamily: C4-dicarboxylate carrier protein

Query Match 64.0%; Score 16; DB 2; Length 98;
 Best Local Similarity 22.2%; Pred. No. 1.2e+02;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
 DB 8 CSAETRAAC 16

RESULT 49
 S76421
 ferredoxin [2Fe-2S] - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C:Accession: S76421
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
 s.
 A:Reference number: S74322; MUID:97061201; PMID:8905231
 A:Accession: S76421
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-103 <KAN>
 A:Cross-references: UNIPROT:P74449; UNIPARC:UPI00000D7139; EMBL:D90915; GB:AB001339; NID
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Keywords: 2Fe-2S; iron-sulfur protein; metalloprotein
 F:28-84/Domain: ferredoxin [2Fe-2S] homology <FER>
 F:43,48,51,83/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 64.0%; Score 16; DB 2; Length 103;
 Best Local Similarity 22.2%; Pred. No. 1.2e+02;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
 DB 83 CAAYATSDC 91

RESULT 50
 J00863
 hypothetical 11.6k protein - Escherichia coli retron Ec67
 C:Species: Escherichia coli retron Ec67
 C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
 C:Accession: J00863
 R:Hsu, M.Y.; Inouye, M.; Inouye, S.
 Proc. Natl. Acad. Sci. U.S.A. 87, 9454-9458, 1990
 A:Title: Retron for the 67-base multicopy single-stranded DNA from Escherichia coli: a p
 A:Reference number: J00851; MUID:91067724; PMID:1701261
 A:Accession: J00863
 A:Molecule type: DNA
 A:Residues: 1-104 <HSU>
 A:Cross-references: UNIPROT:P21322; UNIPARC:UPI000013BF85; GB:M55249; NID:g145143; PIDN:
 A:Experimental source: E. coli strain Cl-1
 C:genetics:
 A:Note: insertion site is equivalent to 19 min of E. coli K12 genetic map
 C:Superfamily: Escherichia coli retron Ec67 hypothetical 11.6k protein

Query Match 64.0%; Score 16; DB 2; Length 104;
 Best Local Similarity 22.2%; Pred. No. 1.2e+02;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
 DB 5 CSAESAHC 13

Search completed: January 4, 2006, 16:10:27
 Job time : 10.313 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 ; Search time 58.7739 Seconds
(without alignments)
108.037 Million cell updates/sec

Title: US-09-932-322-10
Perfect score: 25
Sequence: 1 CXXXXXXC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : Uniprot_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	72.0	144	2	061HG6_DROME
2	18	72.0	155	2	P91214_CAEEL
3	18	72.0	269	2	07PR07_ANOZA
4	18	72.0	281	2	08GQ05_MAIZE
5	18	72.0	281	2	08LKK0_MAIZE
6	18	72.0	347	2	08Q550_9BETA
7	18	72.0	416	2	09ST87_ORYSA
8	18	72.0	468	2	07XST8_ORYSA
9	18	72.0	470	2	09LFB4_ARATH
10	18	72.0	504	2	09XZ06_LEIMA
11	18	72.0	510	2	08S4F6_ARATH
12	18	72.0	513	2	07ZX21_XENIA
13	18	72.0	518	2	0640H3_XENIA
14	18	72.0	569	2	07OKT3_GIALA
15	18	72.0	684	2	07XV79_ORYSA
16	18	72.0	772	1	DILA_BRARE
17	18	72.0	828	2	04WY29_ASPEU
18	18	72.0	1013	2	04SR96_TETNG
19	18	72.0	1031	2	04QFB5_LEIMA
20	18	72.0	1107	2	05CV70_CRYPV
21	18	72.0	1199	1	E75BC_DROME
22	18	72.0	1355	1	E75BA_DROME
23	18	72.0	1365	1	09VU79_DROME
24	18	72.0	1412	1	E75B8_DROME
25	18	72.0	1793	2	09W451_DROME
26	18	72.0	1836	2	08MP07_DROME
27	18	72.0	1894	2	04OS15_LEIMA
28	18	72.0	2087	2	04S488_TETNG
29	17	68.0	50	2	09L381_VIRCH
30	17	68.0	71	2	04RA29_TETNG
31	17	68.0	96	2	08VD96_MESAU

32	17	68.0	108	2	04P4H6_USTMA
33	17	68.0	108	2	09ZSP6_ARATH
34	17	68.0	115	2	07ZD14_DESVH
35	17	68.0	117	2	06VZR1_CNPV
36	17	68.0	117	2	04TWL2_ASP
37	17	68.0	119	2	04TWL3_ASP
38	17	68.0	119	2	04TWL5_ASP
39	17	68.0	119	2	04TWL6_ASP
40	17	68.0	119	2	04TWL7_ASP
41	17	68.0	119	2	04TWL8_ASP
42	17	68.0	119	2	04TWL9_ASP
43	17	68.0	119	2	04TWL0_ASP
44	17	68.0	125	2	05YPN5_NOCFA
45	17	68.0	133	2	04RT14_TETNG
46	17	68.0	134	2	06ZRF5_HUMAN
47	17	68.0	147	2	09CMW3_MOUSE
48	17	68.0	155	2	06F340_ORYSA
49	17	68.0	155	2	04TWL4_ASP
50	17	68.0	157	1	C1025_MOUSE
51	17	68.0	157	1	C1025_MOUSE
52	17	68.0	157	2	04Q6M1_LEIMA
53	17	68.0	160	2	04L116_AZOVI
54	17	68.0	163	2	09BIM0_GIALA
55	17	68.0	166	2	05T594_HUMAN
56	17	68.0	168	2	06K5J0_ORYSA
57	17	68.0	170	2	08PRC3_XANAC
58	17	68.0	171	2	04L182_9ENTR
59	17	68.0	171	2	09JGN6_SALTY
60	17	68.0	172	2	05T590_HUMAN
61	17	68.0	172	2	06K2N4_ORYSA
62	17	68.0	175	2	09VSQ7_DROME
63	17	68.0	183	2	05T595_HUMAN
64	17	68.0	186	1	C1025_BRARE
65	17	68.0	191	2	05Q2H7_BRARE
66	17	68.0	199	2	076588_CABBL
67	17	68.0	202	2	06EXY9_CANGA
68	17	68.0	204	2	06YU14_ORYSA
69	17	68.0	207	2	08N1Y5_HUMAN
70	17	68.0	212	1	YH87_YEAST
71	17	68.0	214	2	0538F6_MYCTU
72	17	68.0	214	2	07U158_MYCTU
73	17	68.0	222	2	04V5Y8_DROME
74	17	68.0	225	2	09LJBT_ARATH
75	17	68.0	228	2	04V5S4_DROME
76	17	68.0	244	2	09Y9Q2_AERYE
77	17	68.0	247	2	06BP27_DEBDA
78	17	68.0	248	2	09US89_SCHPO
79	17	68.0	254	2	06WMS9_BRARE
80	17	68.0	257	2	06ZDL1_ORYSA
81	17	68.0	259	2	083212_TREPA
82	17	68.0	262	1	W0X3A_MAIZE
83	17	68.0	265	1	W0X3B_MAIZE
84	17	68.0	273	2	075H84_ORYSA
85	17	68.0	281	2	057N21_SALGH
86	17	68.0	281	2	05PHQ3_SALPA
87	17	68.0	281	2	08Z796_SALTY
88	17	68.0	281	2	08ZP77_SALTY
89	17	68.0	283	2	07D973_MYCTU
90	17	68.0	286	2	06GQ06_XENIA
91	17	68.0	290	2	06A411_PROAC
92	17	68.0	291	2	08W318_PROBI
93	17	68.0	293	1	XTH31_ARATH
94	17	68.0	298	2	082599_ARATH
95	17	68.0	305	2	04LMH9_9RUTK
96	17	68.0	305	2	067L98_SYWTH
97	17	68.0	306	2	07F8S0_ORYSA
98	17	68.0	306	2	06ZJ81_BURMA
99	17	68.0	308	2	05GM13_CAEEL
100	17	68.0	309	2	08GY34_ARATH
101	17	68.0	309	2	04NURS_9BELT
102	17	68.0	310	2	05N9V7_ORYSA
103	17	68.0	311	2	06EP90_ORYSA
104	17	68.0	311	2	0635X3_BURPS

04P4H6	usutilago ma
09ZSP6	arabidopsis
07ZD14	desulfovibrio
06VZR1	canarypox v
04TWL2	affrican swi
04TWL3	affrican swi
04TWL5	affrican swi
04TWL6	affrican swi
04TWL7	affrican swi
04TWL8	affrican swi
04TWL9	affrican swi
04TWL0	affrican swi
05YPN5	nocardia fa
04RT14	tetradon n
06ZRF5	mus musculus
09CMW3	mus musculus
06F340	oryza sativ
04TWL4	affrican swi
C1025	mus musculus
C1025	mus musculus
04Q6M1	leishmania
04L116	azotobacter
09BIM0	giardia lam
05T594	homo sapien
06K5J0	oryza sativ
08PRC3	xanthomonas
04L182	salmoneilla
09JGN6	salmoneilla
05T590	homo sapien
06K2N4	oryza sativ
09VSQ7	drosophila
05T595	homo sapien
06EXY9	candida gla
06YU14	oryza sativ
08N1Y5	homo sapien
YH87	saccharomyc
0538F6	mycobacteri
07U158	mycobacteri
04V5Y8	drosophila
09LJBT	arabidopsis
04V5S4	drosophila
09Y9Q2	deoxyryum p
06BP27	deoxyryum p
09US89	echinosach
06WMS9	branchiosteo
06ZDL1	oryza sativ
083212	treponema p
W0X3A	zea mays (m
W0X3B	zea mays (m
075H84	oryza sativ
057N21	salmoneilla
05PHQ3	salmoneilla
08Z796	salmoneilla
08ZP77	salmoneilla
07D973	mycobacteri
06GQ06	xenopus lae
06A411	proionidac
08W318	vitis labru
P93014	arabidopsis
082599	arabidopsis
04LMH9	burkholderi
067L98	syndiotacte
07F8S0	oryza sativ
06ZJ81	burkholderi
05GM13	caenorhabdi
08GY34	arabidopsis
04NURS	anaeromyxob
05N9V7	oryza sativ
06EP90	oryza sativ
0635X3	burkholderi


```
981 16 64.0 571 2 Q6QH6_MOUSE Q6QH6_MOUSE
982 16 64.0 571 2 Q5EG07_ICPU Q5EG07_ICPU
983 16 64.0 572 2 Q5IB66_9NICT Q5IB66_spodopora
984 16 64.0 574 2 Q828T3_STRAW Q828T3_streptomyc
985 16 64.0 575 2 Q7QNV8_GITAL Q7QNV8_giardia lam
986 16 64.0 575 2 Q7XNV7_ORISA Q7XNV7_oryza sativ
987 16 64.0 577 2 Q7UX93_RHOBA Q7UX93_rhodospirill
988 16 64.0 578 2 Q7SG68_NEUCR Q7SG68_neutrospora
989 16 64.0 581 2 Q96MM9_BOTCI Q96MM9_botrytis ci
990 16 64.0 582 2 Q5JP29_ORYSA Q5JP29_oryza sativ
991 16 64.0 587 1 SEL10_CABEL Q93794_caenorhabdi
992 16 64.0 588 2 Q6GUE9_TOBAC Q6GUE9_nicotiana t
993 16 64.0 588 2 Q5LXK6_BACPN Q5LXK6_bacteroides
994 16 64.0 588 2 Q64XS9_BACPR Q64XS9_bacteroides
995 16 64.0 592 2 Q60018_PICAN Q60018_pichia anqu
996 16 64.0 592 2 Q98H21_RHITO Q98H21_rhizobium l
997 16 64.0 595 2 Q9C2K3_NEUCR Q9C2K3_neutrospora
998 16 64.0 595 2 Q5VXL7_HUMAN Q5VXL7_homo sapien
999 16 64.0 596 1 HMDH1_SOLITU P48020_solanium tub
1000 16 64.0 596 2 Q4K8Z7_PSEPS Q4K8Z7_pseudomonas
```

ALIGNMENTS

RESULT 1
Q6IKH6_DROME PRELIMINARY; PRT; 144 AA.

```
ID Q6IKH6_DROME PRELIMINARY; PRT; 144 AA.
AC Q6IKH6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE HDCl2427.
GN ORFNames=HDC12427;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14709175; DOI=10.1186/gb-2003-5-1-r3;
RA Hild M., Beckmann B., Haas S.A., Koch B., Solovyev V., Busold C.,
RA Fellenberg K., Boutros M., Vingron M., Sauer F., Hohenfeld J.D.,
RA Paro R.;
RT "An integrated gene annotation and transcriptional profiling approach
RT towards the full gene content of the Drosophila genome."
RC Genome Biol. 5:RESEARCH0003.1-RESEARCH0003.17(2003).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL: BK002390; DAA03896.1; -; Genomic DNA.
SQ SEQUENCE 144 AA; 15277 MW; 1E4588B8C183D5F2 CRC64;
```

Query Match 72.0%; Score 18; DB 2; Length 144;
Best Local Similarity 22.2%; Pred. No. 27;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
DB 37 CATTASASC 45

RESULT 2
P91214_CABEL PRELIMINARY; PRT; 155 AA.

```
ID P91214_CABEL PRELIMINARY; PRT; 155 AA.
AC P91214;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE C-type lectin protein 45.
GN Name=clec-45; ORFNames=F07C4.2;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
```

OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1];
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2; PubMed=9851916;
RX MEDLINE=99069613; Pubmed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).

```
DR EMBL: U80023; AAC24036.1; -; Genomic_DNA.
DR PIR: T28945; T28945.
DR Ensembl: F07C4.2; Caenorhabditis elegans.
DR WormBase: WBGene0017199; F07C4.2.
DR WormPep: F07C4.2; C809201.
DR GO: GO:0005529; F-sugar binding; IEA.
DR InterPro: IPR01304; Lectin_C.
DR SMART: PF00059; Lectin_C; 1.
DR PROSITE: PS00615; C TYPE LECTIN 1; UNKNOWN_1.
DR PROSITE: PS50041; C TYPE LECTIN_2; 1.
KM Complete proteome; Lectin.
SQ SEQUENCE 155 AA; 16285 MW; 956F04FDCB0F93 CRC64;
```

Query Match 72.0%; Score 18; DB 2; Length 155;
Best Local Similarity 22.2%; Pred. No. 28;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
DB 92 CTTASSSSC 100

RESULT 3

```
ID Q7PR07_ANOGA PRELIMINARY; PRT; 269 AA.
AC Q7PR07;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE ENSANGP0000001657 (Fragment).
GN ORFNames=ENSANGG00000001387;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxId=180454;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;  
RT "Anopheles gambiae re-annotation.";  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2];
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;  
CC Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAB01008847; BAA06779.3; -; Genomic_DNA.
FT NON_TER 1 1
FT NON_TER 269 269
SQ SEQUENCE 269 AA; 24228 MW; 25BBFF71FD71F1F2 CRC64;
```

Query Match 72.0%; Score 18; DB 2; Length 269;
Best Local Similarity 22.2%; Pred. No. 34;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
DB 228 CTSSTSSSC 236

RESULT 4

O6GZ05 MAIZE

ID O6GZ05 MAIZE PRELIMINARY; PRT; 281 AA.

AC O6GZ05; 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE DREB-like protein.

GN Name=drebl;

OS Zea mays (Maize).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC PACCAD clade; Panicoideae; Andropogoneae; Zea.

NCBI_TaxID=4577;

OX NCBI_TaxID=4577;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Qiu F., Li J., Zhao J., Chen S.-Y., Liu O.;

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF48789; AAN7673.1; -; mRNA.

DR HSSP: O80337; 2GCC.

DR GO: GO:0005634; C:nucleus; IEA.

DR GO: GO:0003700; F:transcription factor activity; IEA.

DR InterPro: IPR001471; TF_ERF.

DR Pfam: PF00847; AP2; 1.

DR PRINTS: PR00367; ETHRSPELEMT.

DR PRODOM: PD001423; TF_ERF; 1.

DR SMART: SM00380; AP2; 1.

SQ SEQUENCE 281 AA; 29346 MW; A0902A7FE0AB978 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 281;

Best Local Similarity 22.2%; Pred. No. 35;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9

DB 11 CSATTTSSC 19

RESULT 5

ID OBLKX0 MAIZE PRELIMINARY; PRT; 281 AA.

AC OBLKX0;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE DRE binding factor 2.

GN Name=dre2;

OS Zea mays (Maize).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC PACCAD clade; Panicoideae; Andropogoneae; Zea.

NCBI_TaxID=4577;

OX NCBI_TaxID=4577;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA MEDLINE=22058745; PubMed=12061899;

RC Kizis D., Pages M.;

RT "Maize DRE-binding proteins DBF1 and DBF2 are involved in rab17

RT regulation through the drought-responsive element in an ABA-dependent

RT pathway.";

RL Plant J. 30:679-689 (2002).

DR EMBL: AF493799; AAM80485.1; -; mRNA.

DR HSSP: O80337; 2GCC.

DR GO: GO:0005634; C:nucleus; IEA.

DR GO: GO:0003700; F:transcription factor activity; IEA.

DR InterPro: IPR001471; TF_ERF.

DR Pfam: PF00847; AP2; 1.

DR PRINTS: PR00367; ETHRSPELEMT.

DR PRODOM: PD001423; TF_ERF; 1.

DR SMART: SM00380; AP2; 1.

SQ SEQUENCE 281 AA; 29332 MW; CBEO2A7FE0AB978 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 281;

Best Local Similarity 22.2%; Pred. No. 35;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9

DB 11 CSATTTSSC 19

RESULT 6

ID O80S50 9BETA PRELIMINARY; PRT; 347 AA.

AC O80S50;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE UL38.

OS Pongine herpesvirus 4 (Chimpanzee cytomegalovirus).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Betaherpesvirinae; Cytomegalovirus.

NCBI_TaxID=188763;

OX NCBI_TaxID=188763;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=22421467; PubMed=12533697; DOI=10.1099/vir.0.18606-0;

RA Davidson A.J., Dolan A., Akter P., Addison C., Dargan D.J.,

RT Alcendor D.J., McGeoch D.J., Hayward G.S.;

RT "The human cytomegalovirus genome revisited: comparison with the

RT chimpanzee cytomegalovirus genome."

RT J. Gen. Virol. 84:17-28 (2003).

DR EMBL: AF480884; AAM00688.1; -; Genomic DNA.

SQ SEQUENCE 347 AA; 39141 MW; 34B18107D3501982 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 347;

Best Local Similarity 22.2%; Pred. No. 38;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9

DB 312 CSSTSSSTC 320

RESULT 7

ID O9ST87 ORYZA PRELIMINARY; PRT; 416 AA.

AC O9ST87;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE CA303710.1 protein.

GN Name=q3037.10;

OS Oryza sativa (rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Eriharoidae; Oryzaeae; Oryza.

NCBI_TaxID=4530;

OX NCBI_TaxID=4530;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=DNA.

RA Hong G., Chen Z.;

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ245900; CAB53483.1; -; Genomic DNA.

DR Gramene: O9ST87; -.

DR GO: GO:0004629; F:phospholipase C activity; IEA.

DR GO: GO:0007242; P:intracellular signaling cascade; IEA.

DR InterPro: IPR000909; PL_PLC_X.

DR PROSITE: PS50007; PIP2C_X DOMAIN; 1.

SQ SEQUENCE 416 AA; 45736 MW; 332A261949FBE6D CRC64;

Query Match 72.0%; Score 18; DB 2; Length 416;

Best Local Similarity 22.2%; Pred. No. 41;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 34 CTASASASC 42

RESULT 8
Q9XST8 ORYA
AC Q9XST8-ORYA PRELIMINARY; PRT; 468 AA.

DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DE OSJBA0039K24.18 protein.

OS Name=OSJBA0039K24.18;
Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RX PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Meng Q., Zhang L., Lu Y., Mu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Liu X., Lu T., Li C., Wu Y., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.,
RT "Sequence and analysis of rice chromosome 4."
RL Nature 420:316-320(2002).
DR EMBL; AL066637; CAB01799.2; -; Genomic_DNA.
DR Gramene; Q9XST8; -;
DR GO; GO:0004629; P:phospholipase C activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR000909; PI_PLC_X.
DR PROSITE; PS50007; PIP2C_X_DOMAIN; 1.
SQ SEQUENCE 468 AA; 48671 MW; 49D661AA5338A33 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 468;
Best Local Similarity 22.2%; Pred. No. 43;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 34 CTASASASC 42

RESULT 9
Q9LFB4 ARATH
ID Q9LFB4 ARATH PRELIMINARY; PRT; 470 AA.

AC Q9LFB4;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypochemical protein F7J8_200.

GN Name=F7J8_200;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RA Bevan M., Zimmermann W., Gruenisen A., Wambutt R., Kalicki J.,
RA Wohlmann P., Smith A., Bancroft I., Mewes H.W., Rudd S., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
[2]

RP NUCLEOTIDE SEQUENCE.
RU EU Arabidopsis sequencing project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL131189; CAB69850.1; -; Genomic_DNA.
DR PIR; T45962; T45962.

DR GO; GO:0009058; P:biogenesis; IEA.
DR InterPro; IPR001296; Glyco_trans_1.
DR Pfam; PF00534; Glyco_transf_1; I.
KW Hypochemical protein.
SQ SEQUENCE 470 AA; 51797 MW; A7CBB8DA8EF5B6 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 470;
Best Local Similarity 22.2%; Pred. No. 43;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 23 CSSSASTSC 31

RESULT 10
Q9XZ8 LEIMA
ID Q9XZ8 LEIMA PRELIMINARY; PRT; 504 AA.

AC Q9XZ8;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Possible surface antigen.

GN Name=L302.01;
OS Leishmania major.
OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome."
RL Genome Res. 8:135-145(1998).

RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Ivens A.C., Murphy L., Quail M., Harris D., Oliver K., Lawson D.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL389894; CAC22669.1; -; Genomic_DNA.
DR InterPro; IPR006058; 2Fe2S_Fd_BS.
DR InterPro; IPR006212; Furin_repeat.
DR InterPro; IPR002049; Laminin_EGF.
DR SMART; SM00181; EGF; 3.

DR PROSITE; PS00197; 2FE2S_FERREDOXIN; UNKNOWN 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; UNKNOWN 1.
SQ SEQUENCE 504 AA; 52823 MW; B75450B5FE599263 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 504;
Best Local Similarity 22.2%; Pred. No. 44;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 161 CASTTASTC 169

RESULT 11
Q8S4F6 ARATH
ID Q8S4F6 ARATH PRELIMINARY; PRT; 510 AA.

AC Q8S4F6;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)

DE Sulfolipid synthase (Hypothetical protein At5g01220).
GN Name=SOD2; Synonyms=AC5901220;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=11960029; DOI=10.1073/pnas.082696499;
RT "Arabidopsis disrupted in SOD2 encoding sulfolipid synthase is
impaired in phosphate-limited growth."
RL Proc. Natl. Acad. Sci. U.S.A. 99:5732-5737(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
Onodera C.S., Quach H.L., Tang C.C., Toriumi M., Wong C., Wu H.C.,
Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,
Palm C.J., Shim P., Southwick A., Tripp M.G., Wu T., Davis R.W.,
R. Becker J.R., Theologis A.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: ARA54354; AM18913.1; -; mRNA.
DR EMBL: BT005796; AAO64198.1; -; mRNA.
DR GO: GO:0009058; P: biosynthesis; IEA.
DR InterPro: IPR001296; Glyco trans 1.
DR Pfam: PF00534; Glycos_transf_1; 1.
DR Hypothetical protein_510 AA; 56630 MW; 20E57D318DBED68 CRC64;
SQ SEQUENCE

Query Match 72.0%; Score 18; DB 2; Length 510;
Best Local Similarity 22.2%; Pred. No. 44;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 23 CSSSATSC 31
RESULT 12
ID 07X21_XENLA PRELIMINARY; PRT; 513 AA.
AC 07X21;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE MGC3520 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Rahs S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Vallion D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
R. Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smolius D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC046254; AAH46254.1; -; mRNA.
DR InterPro: IPR006209; EGF-like.
DR InterPro: IPR003323; OTU.
DR Pfam: PF02338; OTU; 1.
DR PROSITE: PS00022; EGF 1; UNKNOWN_1.
DR PROSITE: PS00802; OTU; 1.
SQ SEQUENCE 513 AA; 55868 MW; A8734AD686FEEDA CRC64;

Query Match 72.0%; Score 18; DB 2; Length 513;
Best Local Similarity 22.2%; Pred. No. 44;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 376 CSSATMAAC 384

RESULT 13
ID 0640H3_XENLA PRELIMINARY; PRT; 518 AA.
AC 0640H3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE LOC494714 protein.
GN Name=LOC494714;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Rahs S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Vallion D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.M., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Kidney;
 RA Klein S., Gerhard D.S.,
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC082654; AA082654.1; -; mRNA.
 DR InterPro: IPR003323; OTU.
 DR Pfam: PF02338; OTU; 1.
 DR PROSITE: PS50802; OTU; 1.
 SQ SEQUENCE 518 AA; 56209 MW; 2099AB07C6C93EC1 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 518;
 Best Local Similarity 22.2%; Pred. No. 45;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXX 9
 Db 381 CSSTAAAC 389

RESULT 14
 Q70XT3 GIALA PRELIMINARY; PRT; 569 AA.
 AC Q70XT3;
 DT 01-MAR-2004 (TREMBLrel. 26, Created)
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
 DE GIP 77 40692 38983.
 OS Giardia lamblia ATCC 50803.
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinales; Giardia.
 NC NBI_TaxID=184922;
 RN [1]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=WB C6;
 RX MEDLINE=21167845; PubMed=11104758; DOI=10.1074/jbc.M006589200;
 RA Abel E.S., Davies B.J., Robles L.D., Loflin C.E., Gillin F.D.,
 RA Chakrabarti R.,
 RT "Possible roles of protein kinase A in cell motility and excystation
 RT of the early diverging eukaryote Giardia lamblia.";
 RL J. Biol. Chem. 276:10320-10329(2001).
 RN [2]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=WB C6;
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
 RA Olsen G.J., Sogin M.L.,
 RT "Draft sequence of the Giardia lamblia genome.";
 RT Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL: AACB01000056; EAA39878.1; -; Genomic DNA.
 DR GO: GO:0005489; F:electron transporter activity; IEA.
 DR GO: GO:0005506; F:iron ion binding; IEA.
 DR GO: GO:0006118; P:electron transport; IEA.
 DR InterPro: IPR001450; 4Fe4S Fe-S bd.
 DR InterPro: IPR005127; Giardia_VSF.
 DR Pfam: PF03302; VSP; 1.
 DR PRINTS: PR00353; 4FE4SFRDOXIN.
 KM Electron transport; Transport.
 SQ SEQUENCE 569 AA; 59086 MW; EB0DF4E2B8CC3F20 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 569;
 Best Local Similarity 22.2%; Pred. No. 46;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXX 9
 Db 464 CSSTAAAC 472

RESULT 15
 Q7XV79 ORYZA PRELIMINARY; PRT; 684 AA.
 AC Q7XV79;
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
 DE OSJNB0076A22.6 protein.
 GN Name=OSJNB0076A22.6;
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 NC NBI_TaxID=4530;
 RN [1]
 RN NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22337377; PubMed=12447439; DOI=10.1038/nature01183;
 RA Feng Q., Zhang Y., Hao P., Wang S., Pu G., Huang Y., Li Y., Zhu J.,
 RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
 RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
 RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Gan J.,
 RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Jin H.,
 RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
 RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
 RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
 RA Lan L., Lai Y., Cheng Z., Gu W., Jiang D., Li J., Hong G., Xue Y.,
 RT "Sequence and analysis of rice chromosome 4.";
 RL Nature 420:316-320(2002).
 RX EMBL: AL663016; CAD40795.3; -; Genomic DNA.
 DR Gramine; Q7XV79;
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO: GO:0016740; F:transferase activity; IEA.
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR008271; Ser_thr_pkin_AS.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00069; Pkinase; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00219; Tyrc; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN 1.
 SQ SEQUENCE 684 AA; 75877 MW; 8CB1910F10409B14 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 684;
 Best Local Similarity 22.2%; Pred. No. 50;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXX 9
 Db 20 CSAAATATC 28

RESULT 16
 DLA-BRARE STANDARD; PRT; 772 AA.
 ID DLA-BRARE
 AC Q6DI18; O57462;
 DT 13-SEP-2005 (Rel. 48, Created)
 DT 13-SEP-2005 (Rel. 48, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Delta-like protein A precursor (DeltaA protein).
 GN Name=dlp;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 CC Cyprinidae; Danio.
 CC NCBI_TaxID=7955;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA], TISSUE SPECIFICITY, AND DEVELOPMENTAL
 RP STAGE.
 RX MEDLINE=98165392; PubMed=9425133;
 RA Appel B., Eiseen J.S.;
 RT "Regulation of neuronal specification in the zebrafish spinal cord by
 RT Delta function.";
 RL Development 125:371-380 (1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Embryo;
 RG NIH - Zebrafish Gene Collection (ZGC) project;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP FUNCTION.
 RX PubMed=10572043;
 RA Riley B.B., Chiang M.-Y., Farmer L., Heck R.;
 RT "The delta gene of zebrafish mediates lateral inhibition of hair
 RT cells in the inner ear and is regulated by pax2.1.";
 RL Development 126:5669-5678 (1999).
 RN [4]
 RP INTERACTION WITH MIB, AND UBIQUITINATION.
 RX PubMed=15013799; DOI=10.1016/j.ydbio.2003.11.010;
 RA Chen W., Corliss D.C.;
 RT "Three modules of zebrafish Mind bomb work cooperatively to promote
 RT Delta ubiquitination and endocytosis.";
 RL Dev. Biol. 267:361-373 (2004).
 CC -1- FUNCTION: Acts as a ligand for Notch receptors and is involved in
 CC primary neurogenesis. Can activate Notch receptors, thereby
 CC playing a key role in lateral inhibition, a process that prevents
 CC the immediate neighbors of each nascent neural cell from
 CC simultaneously embarking on neural differentiation. In inner ear,
 CC it prevents adjacent cells from adopting the same cell fate.
 CC -1- SUBUNIT: Interacts with mib.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -1- TISSUE SPECIFICITY: Expressed in nervous system. In the developing
 CC nervous system, it is expressed in overlapping regions with deltab
 CC (dlb) and deltal (dl), but differs the neural plate: it is
 CC expressed in patches of contiguous cells with dlid, while dila is
 CC apparently confined to the scattered cells within those patches
 CC that differentiate as neurons. Expressed in hair cells of inner
 CC ear.
 CC -1- DEVELOPMENTAL STAGE: Initiated in the neuroectoderm before that of
 CC dld. In the developing trunk neural plate and neural tube, it is
 CC initiated in the epiblast prior to completion of gastrulation. At
 CC the 2- to 3-somite stage (10.5 hours) low levels are distributed
 CC throughout the trunk CNS, with cells expressing higher levels
 CC found in the medial and lateral regions of the neural plate. These
 CC regions correspond to the positions at which primary motoneurons
 CC and Rohon Bead neurons (Rbns) originate. Cells expressing high
 CC levels do not form contiguous domains. Rather, single cells or
 CC small clusters of several cells showing high expression are
 CC interspersed with cells having lower expression. Expressed is
 CC specific to the developing nervous system, and continues to be
 CC expressed broadly in the CNS throughout neurogenesis. Expressed in
 CC cells specified for neuronal fates. At 24 hours, and throughout
 CC later embryogenesis, it is broadly expressed in the spinal cord,
 CC suggesting that it is expressed by many types of cells. Expressed
 CC as neuronal specification occurs and is subsequently down-
 CC regulated in cells that have acquired specific neuronal fates.
 CC -1- PTM: Ubiquitinated by mib, leading to its endocytosis and
 CC subsequent degradation.
 CC -1- SIMILARITY: Contains 1 DSL domain.
 CC -1- SIMILARITY: Contains 8 EGF-like domains.
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 726.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL: AF030031; AAC41249.1; ALT_SRO; mRNA.
 CC EMBL: BC075742; AAH75742.1; -, mRNA.
 CC ZFIN: ZDB-GENE-980526-29; dila.
 CC InterPro: IPR000152; Asx_hydroxyl_S.
 CC InterPro: IPR001774; DSL.
 CC InterPro: IPR000742; EGF_2.
 CC InterPro: IPR001881; EGF_Ca_bd.
 CC InterPro: IPR001438; EGF_II.
 CC InterPro: IPR006209; EGF_like.
 CC InterPro: IPR006210; IEGF.
 CC InterPro: IPR016511; MNML.
 CC Pfam: PF01414; DSL_1.
 CC Pfam: PR00008; EGF_6.
 CC Pfam: PF07657; MNML_1.
 CC PRINTS: PR00010; EGFBLD.
 CC SMART: SM00051; DSL_1.
 CC SMART: SM00181; EGF_8.
 CC SMART: SM00179; EGF_CA; 6.
 CC PROSITE: PS00010; ASX_HYDROXYL; 2.
 CC PROSITE: PS1051; DSL_1.
 CC PROSITE: PS00022; EGF_1; 8.
 CC PROSITE: PS01186; EGF_2; 8.
 CC PROSITE: PS50026; EGF_3; 7.
 CC PROSITE: PS01187; EGF_CA; 1.
 CC Calcium: Developmental protein; Differentiation; EGF-like domain;
 CC Glycoprotein; Neurogenesis; Notch signaling pathway; Repeat; Signal;
 CC Transmembrane; Ubl conjugation.
 CC KEGG: CHAIN 1
 CC KEGG: TOPO_DOM 21
 CC KEGG: TRANSMEM 537
 CC KEGG: TOPO_DOM 558
 CC KEGG: DOMAIN 179
 CC KEGG: DOMAIN 225
 CC KEGG: DOMAIN 257
 CC KEGG: DOMAIN 290
 CC KEGG: DOMAIN 330
 CC KEGG: DOMAIN 366
 CC KEGG: DOMAIN 407
 CC KEGG: DOMAIN 445
 CC KEGG: DOMAIN 483
 CC KEGG: COMBIAS 697
 CC KEGG: CARBOHYD 479
 CC KEGG: DISULFID 228
 CC KEGG: DISULFID 232
 CC KEGG: DISULFID 259
 CC KEGG: DISULFID 265
 CC KEGG: DISULFID 278
 CC KEGG: DISULFID 294
 CC KEGG: DISULFID 300
 CC KEGG: DISULFID 318
 CC KEGG: DISULFID 334
 CC KEGG: DISULFID 339
 CC KEGG: DISULFID 356
 CC KEGG: DISULFID 372
 CC KEGG: DISULFID 377
 CC KEGG: DISULFID 395
 CC KEGG: DISULFID 411
 CC KEGG: DISULFID 416
 CC KEGG: DISULFID 433
 CC KEGG: DISULFID 449
 CC KEGG: DISULFID 454
 CC KEGG: DISULFID 471
 CC KEGG: DISULFID 487
 CC KEGG: DISULFID 492
 CC KEGG: DISULFID 509
 CC KEGG: SEQUENCE 772 AA; 84969 MW; 716A01415893576 CRC64;
 CC -----
 CC Query Match 72.0%; Score 18; DB 1; Length 772;
 CC Best Local Similarity 22.2%; Pred. No. 52;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

DB 740 CASASTSAC 748

RESULT 17

Q4WY29 ASPFU PRELIMINARY; PRT; 828 AA.

AC Q4WY29; 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Repressor protein.

GN ORFNames=AF13514830;

OS Aspergillus fumigatus Af293.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

OX NCBI_TaxId=330879;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=AF293;

RA Niernan W., Pain A., Anderson M.J., Wortman J., Kim H., Stanley J.,

RA Arioye J., Berriman M., Abe K., Archer D.B., Bernedo C., Bennett J.,

RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,

RA Farman M., Fedorova N., Fedorova N., Feldlyum T.V., Fischer R.,

RA Foster N., Fraser A., Garcia J.L., Garcia M.U., Godle A.,

RA Goldman G.H., Goni K., Griffith-Jones S., Gwilliam R., Haas B.,

RA Haas H., Harris D., Horluchi H., Huang J., Humphrey S., Jimenez J.,

RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,

RA Kumagai T., Latton A., Latge J.-P., Li W., Lord A., Lu C.,

RA Majors W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,

RA Mouyia I., Mulligan S., Murphy L., O'Neill S., Paulsen I.,

RA Penativa M.A., Petrea M., Price C., Pritchard B.L., Quail M.A.,

RA Rabinovitch E., Rawlins N., Rajandream M.-A., Reilhard U.,

RA Renaldi H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,

RA Roming C.M., Ruter S., Salzberg S.L., Sanchez M.,

RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,

RA Takeuchi M., Tekala F., Turner G., Vazquez de Alana C.R., Weidman J.,

RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,

RA Machida M., Hall N., Barrell B., Denning D.W.;

RT "Genomic sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus.";

RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.

CC -1- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL; AAHF0100002; EAL92104.1; -; Genomic DNA.

CC DR EMBL; AAHF0100002; EAL92104.1; -; Genomic DNA.

SQ SEQUENCE 828 AA; 92421 MW; 2B1E831C5B4FF0C6 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 828;

Best Local Similarity 22.2%; Pred. No. 54;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

DB 57 CTTTTTTC 65

RESULT 18

Q4SR96 TETNG PRELIMINARY; PRT; 1013 AA.

AC Q4SR96; 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Chromosome 11 SCAP14528, whole genome shotgun sequence.

DE (Fragment).

GN ORFNames=GSTENG00014015001;

OS Tetraodon nigroviridis (Green puffer).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;

OC Tetraodontidae; Tetraodontidae; Tetraodon.

OX NCBI_TaxId=99883;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,

RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,

RA Nicand S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Segreus B.,

RA Dasila C., Salanoubat M., Levy M., Boudet N., Castellano S.,

RA Anthouard V., Jabin C., Castell V., Katinka M., Vacherie B.,

RA Blemont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,

RA Craud C., Duprat S., Brothier P., Contancieu J.P., Gonzy J.,

RA Parra G., Lardier G., Chaple C., McKernan K.J., McEwan P., Bosak S.,

RA Kellis M., Wolf J.N., Guigo R., Zody M.C., Mesirov J.,

RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,

RA Laudet V., Schachter V., Quetier F., Sautin W., Scarpelli C.,

RA Wincker P., Lander E.S., Weissbach J., Roest Crollins H.;

RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype.";

RL Nature 431:946-957(2004).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RG Genoscope, Whitehead Institute Centre for Genome Research;

RL Submitted (Feb-2004) to the EMBL/GenBank/DBJ databases.

CC -1- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL; CAAB01014528; CAP96836.1; -; Genomic DNA.

DR InterPro; IPR000175; Na/atrian_symport.

DR Pfam; PF00209; SNE; 2.

DR PRINTS; PR00176; NANUSMPORT.

DR Prodom; PD000448; Na/atrian_symport; 2.

DR PROSITE; PS00610; NA_NEUROTRAN_SYMPT_1; 1.

DR PROSITE; PS00754; NA_NEUROTRAN_SYMPT_2; UNKNOWN_1.

DR PROSITE; PS0267; NA_NEUROTRAN_SYMPT_3; 2.

KW Symport; Transmembrane; Transport.

FT NON TER 1

FT NON TER 1013

SQ SEQUENCE 1013 AA; 110756 MW; 3D5SCFE7787652PD CRC64;

Query Match 72.0%; Score 18; DB 2; Length 1013;

Best Local Similarity 22.2%; Pred. No. 58;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

DB 728 CSASTSAC 736

RESULT 19

Q4QFB5 LEIMA PRELIMINARY; PRT; 1031 AA.

AC Q4QFB5; 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Hypothetical protein.

GN ORFNames=LmjF15.0700;

OS Leishmania major.

OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OC NCBI_TaxId=5664;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Frledlin;

RA Peacock C.S., Murphy L., Ivans A.C., Berriman M., Blackwell J.,

RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,

RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;

RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.

DR EMBL; CTO05254; CAJ03294.1; -; Genomic DNA.

DE Hypothetical protein.

SQ SEQUENCE 1031 AA; 112673 MW; 65DB334CA5FCBA CRC64;

Query Match 72.0%; Score 18; DB 2; Length 1031;

Best Local Similarity 22.2%; Pred. No. 59;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 9

Db 791 CSSSSSSSAC 799

RESULT 20
OSCV70_CRYPV PRELIMINARY; PRT; 1107 AA.

AC OSCV70;
DT 10-MAY-2005 (TReMBLrel. 30, Created)
DT 10-MAY-2005 (TReMBLrel. 30, Last sequence update)
DE Hypothetical protein.
GN ORFNames=cg48_5210;
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxId=5807;

RP NUCLEOTIDE SEQUENCE.
STRAIN=Iowa type II;
RX PubMed=15044751; DOI=10.1126/science.1094786;
RA Abrahamson M.S., Templeton T.J., Enomoto S., Abrahante J.B., Zhu G.,
RA Lanco C.A., Deng M., Liu C., Widmer G., Tzipori S., Buck G.A., Xu P.,
RA Bankier A.T., Dear P.H., Konfortov B.A., Spriggs H.F., Iyer L.,
RA Anantharaman V., Aravind L., Kapur V.;
RT "Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*.";
RL Science 304:441-445 (2004).
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.

CC EMBL; AAE0100003; EAK89643.1; -; Genomic DNA.
DR InterPro: IPR012283; 6PGD C core.
DR InterPro: IPR011989; ARM-like.
DR InterPro: IPR012292; Globin-related.
DR InterPro: IPR001524; Glyco_Hydro-6.
DR InterPro: IPR012287; Homeodomain-rel.
DR InterPro: IPR001093; IMPH/GMPase.
DR InterPro: IPR008976; PLAT LH2.
DR InterPro: IPR000783; RNAPol_RPB5.
DR InterPro: IPR009032; Vpu_cyt.
DR InterPro: IPR011991; Wing_hlx_DNA_bd.
DR Pfam: PF00478; IMPDH; 1.
KW Hypothetical protein.
SQ SEQUENCE 1107 AA; 127343 MW; 4BAD9AE836ECB2AF CRC64;

Query Match 72.0%; Score 18; DB 2; Length 1107;
Best Local Similarity 22.2%; Pred. No. 60;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 9

Db 185 CSSSSSSSAC 193

RESULT 21
E75BC DROME
ID E75BC DROME STANDARD; PRT; 1199 AA.
AC P1671; O810S1; O810S3;
DT 01-ANG-1990 (Rel. 15, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Ecdysone-induced protein 75B isoforms C/D (E75-A).
GN Name=E75B; Synonyms=NR1D3; ORFNames=CG8127;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;

RN [1]
RC NUCLEOTIDE SEQUENCE [MRNA], AND ALTERNATIVE SPLICING.
RP STRAIN=Cancon-S;
RX MEDLINE=90249727; PubMed=2110921;
RA Segaves W.A., Hognes D.S.;
RT "The E75 ecdysone-inducible gene responsible for the 75B early puff in
RT Drosophila encodes two new members of the steroid receptor
RT superfamily.";
RL Genes Dev. 4:204-219 (1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benoe P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Gload A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lascko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murty D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacible J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein D.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195 (2000).
RN [3]
RP GENOME REANNOTATION, AND ALTERNATIVE SPLICING.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Betencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).
RN [4]
RP FUNCTION, DEVELOPMENTAL STAGE, AND INDUCTION.
RX MEDLINE=94038699; PubMed=8223281;
RA Huet F., Ruiz C., Richards G.;
RT "Puffs and PCR: the in vivo dynamics of early gene expression during
RT ecdysone responses in *Drosophila*.";
RL Development 118:613-627 (1993).
CC -!- FUNCTION: Implicated in the regulation of ecdysone-triggered gene
hierarchies. Probably plays a key role in mediating the regulation

```

CC of the larval molt by 20-OH-ecdysone.
CC -1 SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1 ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=C; Synonyms=E75A;
CC IsoId=P1671-1; Sequence=Displayed;
CC Name=A; Synonyms=E75B;
CC IsoId=P1672-1; Sequence=External;
CC Name=B; Synonyms=E75C;
CC IsoId=P1055-1; Sequence=External;
CC Name=D;
CC IsoId=P1671-2; Sequence=VSP_014915; VSP_014916;
CC Note=No experimental confirmation available;
CC -1 DEVELOPMENTAL STAGE: In mid instar larvae salivary glands, low
CC basal levels are observed in puff stage 1. Levels increase in late
CC larvae from puff stages 3-10, then decrease abruptly at stage 11.
CC In prepupae, isoform C is the predominant form during the
CC transition between puff stages 18-19. At puff stage 1, expression
CC is also present in the gut. By stage 3 it is present in the wing
CC disks, Malpighian tubules and the fat body. At stage 11,
CC expression is only present in the gut and wing disks.
CC -1 INDUCTION: The expression of this protein is developmentally
CC regulated and is correlated with the 20-OH-ecdysone induced
CC activity of puff 75B.
CC -1 SIMILARITY: Belongs to the nuclear hormone receptor family. NR1
CC subfamily.
CC -1 SIMILARITY: Contains 1 nuclear receptor DNA-binding domain.
CC -1 CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 1164.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL database
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL, X51548; CA35923.1; ALT FRAME, mRNA.
DR EMBL, AE003522; AAN11687.1; -; Genomic DNA.
DR EMBL, AE003522; AAN11689.1; -; Genomic DNA.
DR PIR, A34598; A34598.
DR HSSP, P20393; 1GAS.
DR SMR, P1671; 240-323.
DR TRANSFAC, T01367; -.
DR FLYBASE, FBgn0000568; E1p75B.
DR GO, GO:0004879; F.ligand-dependent nuclear receptor activity; NAS.
DR GO, GO:0019890; P.ecdysis (sensu insecta); IMP.
DR GO, GO:0035072; P.ecdysone-mediated induction of salivary gla. . .; NAS.
DR GO, GO:0007553; P.regulation of ecdysteroid metabolism; IMP.
DR InterPro, IPR001628; Hrmn_rcpt_DNA_bd.
DR InterPro, IPR000536; Hrmn_rcpt_lig_bd.
DR InterPro, IPR001723; Stdnrm_receptor.
DR InterPro, IPR000324; VltD_receptor.
DR Pfam, PF00104; Hormone_recep. 1.
DR Pfam, PF00105; zf-C4; 1.
DR PRINTS, PR00398; STRDHORMNER.
DR PRINTS, PR00447; STROIDFINGER.
DR PRINTS, PR00350; VITAMINDER.
DR PRODOM, PD0000035; Znf_C4steroid; 1.
DR SMART, SM00430; HO1I; 1.
DR SMART, SM00399; Znf_C4; 1.
DR PROSITE, PS00031; NUCLEAR_REC_DBD_1; 1.
DR PROSITE, PS01030; NUCLEAR_REC_DBD_2; 1.
DR Alternative splicing: Developmental protein; DNA-binding;
DR Metal-binding; Nuclear protein; Receptor; Transcription;
DR Transcription regulation; Zinc; Zinc-finger.
KW Transcription regulation; Zinc; Zinc-finger.
FT DNA_BIND 242 318 Nuclear receptor.
FT ZN_FING 245 265 NR_C4-type.
FT ZN_FING 282 306 NR_C4-type.
FT REGION 380 590 Ligand-binding (potential).
FT COMPBIAS 139 195 Gln/His-rich.
FT COMPBIAS 721 756 Gln-rich.
FT COMPBIAS 900 1103 Ser-rich.
FT VARSPIC 1 293 Missing (in isoform D).

```

```

FT FT VARSPIC 294 315 /FTId=VSP_014915.
FT FT VARSPIC 294 315 NNRCCYCRILKICIAVGMGRDA -> MGEELPILKILKGN
FT FT VARSPIC 294 315 VNYHNAP (in isoform D).
FT FT VARSPIC 294 315 /FTId=VSP_014916.
SQ SEQUENCE 1199 AA; 128534 MW; E29372F5E29F55 CRC64;
Query Match 72.0%; Score 18; DB 1; Length 1199;
Best Local Similarity 22.2%; Pred. No. 62;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 CXXXXXXX 9
Db 1070 CXXXXXXX 1078
RESULT 22
E75BA_DROME STANDARD; PRT; 1355 AA.
ID E75BA_DROME STANDB; PRT; 1355 AA.
AC P1672; Q810S2;
DT 01-AUG-1990 (Rel. 15, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Ecdysone-induced protein 75B isoform A (E75-B).
GN Name=E1p75B; Synonyms=NR1D3; ORFNames=CG8127;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND ALTERNATIVE SPLICING.
RC STRAIN=Canton-S;
RX MEDLINE=90249727; PubMed=2110921;
RT Segreaves W.A., Hogness D.S.;
RT "The E75 ecdysone-inducible gene responsible for the 75B early puff in
RT Drosophila encodes two new members of the steroid receptor
RT superfamily."
RL Genes Dev. 4:204-219(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sultion G.G., Wortman J.R., Yandell M.D., Zhang L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan R.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-O., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolhakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heitman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,
RA Jalali M., Kalish F., Kapen G.H., Ke Z., Kemisun J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Svatek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

```

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).
 RP GENOME REANNOTATION, AND ALTERNATIVE SPLICING.
 RA MEDLINE=22426069; PubMed=12537572;
 RA Mista S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminck J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.,
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review,"
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 [4]
 RP FUNCTION, DEVELOPMENTAL STAGE, AND INDUCTION.
 RA MEDLINE=94038699; PubMed=8223281;
 RA Hueb F., Ruiz C., Richards G.,
 RT "Puffs and PCR: the in vivo dynamics of early gene expression during
 RT ecdysone responses in *Drosophila*,"
 RL Development 118:613-627(1993).
 CC -!- FUNCTION: Implicated in the regulation of ecdysone-triggered gene
 CC hierarchies. Probably plays a key role in mediating the regulation
 CC of the larval molt by 20-OH-ecdysone.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=A: Synonyms=E75B;
 CC IsoId=PI7672-1; Sequence=Displayed;
 CC Name=C: Synonyms=E75A;
 CC IsoId=PI7671-1; Sequence=External;
 CC Name=B: Synonyms=E75C;
 CC IsoId=PI3055-2; Sequence=External;
 CC Name=D:
 CC IsoId=PI7671-2; Sequence=External;
 CC Note=No experimental confirmation available;
 CC -!- DEVELOPMENTAL STAGE: In mid instar larvae salivary glands, levels
 CC are low during puff stage 1, increase during puff stages 2-4 and
 CC diminish from stage 5 onwards. In prepupae, isoform A is the
 CC predominant form during puff stage 19 and the transition to stage
 CC 20. By stage 3 it is present in the gut, Malpighian tubules and
 CC the fat body, levels persist beyond stage 11.
 CC -!- INDUCTION: The expression of this protein is developmentally
 CC regulated and is correlated with the 20-OH-ecdysone induced
 CC activity of puff 75B.
 CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NRI
 CC subfamily.
 CC -!- SIMILARITY: Contains 1 nuclear receptor DNA-binding domain.
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 1320.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL, X51549; CNA5924.1, ALT_FRAME, mRNA.
 CC EMBL, AE003522; AAN11688.1, -, Genomic DNA.
 CC PIR, B34598; B34598.
 CC HSSP, P20393; 1A6Y.
 CC SMR, P1672; 415-480.
 CC TRANSFAC, T01368; -.
 CC Flybase: FBgn0000568; E1p75B.
 DR GO, GO:0004879; F, ligand-dependent nuclear receptor activity; NAS.
 DR GO, GO:0018990; P, ecdysone (samen insecta); IMP.
 DR GO, GO:0035072; P, ecdysone-mediated induction of salivary gla. .; NAS.

DR GO, GO:0007553; P, regulation of ecdysteroid metabolism; IMP.
 DR InterPro, IPR001628; Hrmr rcpt DNA bd.
 DR InterPro, IPR000536; Hrmr rcpt lig bd.
 DR InterPro, IPR000354; Involucrin_rcpt.
 DR InterPro, IPR003079; ROR receptor.
 DR InterPro, IPR001723; Scdhrm_receptor.
 DR Pfam, PF00104; Hormone recep; 1.
 DR Pfam, PF00904; Involucrin; 1.
 DR Pfam, PF00105; zf-C4; 1.
 DR PRINTS, PRO1293; RORNUCPCPTR.
 DR PRINTS, PRO0398; STRDHOMNER.
 DR PRINTS, PRO0047; STROIDFINGER.
 DR PRODOM, P0000035; Znf C4steroid; 1.
 DR SMART, SM00430; HOL1; 1.
 DR SMART, SM00399; ZNF C4; 1.
 DR PROSITE, PS51030; NUCLEAR_REC_DBD_2; 1;
 KW Alternative splicing; Developmental protein; DNA-binding;
 KW Metal-binding; Nuclear protein; Receptor; Transcription;
 KW Transcription regulation; Zinc; Zinc-finger.
 FT DNA_BIND 384 474
 FT ZN_FING 387 421
 FT ZN_FING 438 457
 FT COMEBIAS 6 324
 FT COMEBIAS 877 912
 FT COMEBIAS 1056 1259
 FT COMEBIAS 206 206
 FT CONFLICT 244 244
 FT CONFLICT 244 244
 SO SEQUENCE 1355 AA; 147171 MW; 6EF19BACEC56225 CRC64;
 Query Match 72.0%; Score 18; DB 1; Length 1355;
 Best Local Similarity 22.2%; Pred. No. 65;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 CXXXXXXC 9
 DB 1226 CXXXXXXC 1234
 RESULT 23
 ID Q9VJ79 DROME PRELIMINARY; PRT; 1365 AA.
 AC Q9VJ79;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE CG10231-PA.
 GN Name=Pdel1; ORFNames=CG10231;
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 OC NCBI_Taxid=7227;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouch H., Brockstein P., Brodtier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Crowley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Dey S., Davis A.P., Dew I., Dietz S.M.,
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Foster C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Maassman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
[2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celinker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Taber P.B., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
[3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celinker S.E.,
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
RT a genomic perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
[4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Mistrz S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochownik S.E.,
RA Smith C.D., Tuzy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bertencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.,
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
[5]
RP NUCLEOTIDE SEQUENCE.
RG Berkeley *Drosophila* Genome Project:
RA Celinker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.,
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[6]
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RT Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL: AEO03659; AAF53675.2; -; Genomic DNA.
DR Ensembl: CG10231; Drosophila melanogaster.
DR FlyBase: FBgn0032686; Pdel1.
DR GO: GO:0004114; F:3',5'-cyclic-nucleotide phosphodiesterase a. . .; IEA.
DR GO: GO:0003824; F:catalytic activity; IEA.
DR GO: GO:0007165; P:signal transduction; IEA.

DR InterPro: IPR003018; GAF.
DR InterPro: IPR003607; Met_phos_hydro.
DR InterPro: IPR002073; PDBase.
DR Pfam: PF01590; GAF; 2.
DR Pfam: PF00233; PDBase; 1.1.
DR PRINTS: PR00387; PDBSTERASE1.
DR SMART: SM00065; GAF; 2.
DR SMART: SM00471; HGC; 1.
DR PROSITE: PS00126; PDBASE 1; 1.
SQ SEQUENCE 1365 AA; 151173 MW; B0562EFPD5012B56 CRC64;
Query Match 72.0%; Score 18; DB 2; Length 1365;
Best Local Similarity 22.2%; Pred. No. 66;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
DB 1138 CDSSTRASSC 1146
QY 1 CXXXXXXXXX 9
DB 1138 CDSSTRASSC 1146
RESULT 24
E75BB DROME STANDARD; PRT; 1412 AA.
ID E75BB DROME STANDARD; PRT; 1412 AA.
AC P13055; Q9VVM9;
DT 01-JAN-1990 (Rel. 13, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Ecdysone-induced protein 75B isoform B (E75-C).
GN Name:E75B; Synonym=NR1D3; ORFNames=C68127;
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
OX NCBI_TaxId=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND ALTERNATIVE SPLICING.
RC STRAIN=Canton-S; TISSUE=Head;
RX MEDLINE=90016778; PubMed=2508058;
RA Feigl G., Gram M., Pongs O.,
RT "A member of the steroid hormone receptor gene family is expressed in
RT the 20-OH-ecdysone inducible puff 75B in *Drosophila melanogaster*.";
RL Nucleic Acids Res. 17:7167-7178(1989).
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Boyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Baillet R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Brottier P.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Gawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delber A., Deng Z., Mayhew A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunker B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostel C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Svirkas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Massaman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of *Drosophila melanogaster*.";
RA Science 287:2185-2195(2000).
RN [3]
RP GENOME REANNOTATION, AND ALTERNATIVE SPLICING.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Belecourt B.R., Cheliker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RA "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review.";
RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [4]
RP FUNCTION, DEVELOPMENTAL STAGE, AND INDUCTION.
RX MEDLINE=94038699; PubMed=8223281;
RA Huet F., Ruiz C., Richards G.;
RA "Puffs and PCR: the in vivo dynamics of early gene expression during
RT ecdysone responses in *Drosophila*.";
RT Development 118:613-627(1993).
CC -1- FUNCTION: Implicated in the regulation of ecdysone-triggered gene
CC hierarchies. Probably plays a key role in mediating the regulation
CC of the larval molt by 20-OH-ecdysone.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=B; Synonyms=E75C;
CC IsoId=PI3055-2; Sequence=Displayed;
CC Name=C; Synonyms=E75A;
CC IsoId=PI7671-1; Sequence=External;
CC Name=A; Synonyms=E75B;
CC IsoId=PI7672-1; Sequence=External;
CC Name=D;
CC IsoId=PI7671-2; Sequence=External;
CC NOTE=No experimental confirmation available;
CC -1- DEVELOPMENTAL STAGE: In mid instar larvae salivary glands. Levels
CC increase during puff stage 1, then remain relatively constant
CC until the premetamorphic pulse of ecdysone at puff stage 5. Levels
CC increase again in late larvae at puff stages 9-10. At puff stage 1
CC expression is also seen in the gut. Levels are low in the gut,
CC Malpighian tubules, fat body and wing disks between stages 1 and
CC 11.
CC -1- INDUCTION: The expression of this protein is developmentally
CC regulated and is correlated with the 20-OH-ecdysone induced
CC activity of puff 75B.
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NRI
CC subfamily.
CC -1- SIMILARITY: Contains 1 nuclear receptor DNA-binding domain.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
CC in positions 15, 543, 576, 593, 614 and 305.
CC -----
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC DR EMBL; X15586; CAA3611.1; ALT_FRAME; mRNA.
CC DR EMBL; AE003522; AAF49282.3; -; Genomic_DNA.
CC DR PIR; S05979; S05979.
CC DR HSSP; P20393; 1GAS.
CC DR SMR; PI3055; 449-532.

DR Ensembl; CG8127; *Drosophila melanogaster*.
DR FlyBase; Fg0000568; Eip75B.
DR GO; GO:0004879; F.1.ligand-dependent nuclear receptor activity; NAS.
DR GO; GO:0018990; P.ecdysis (sensu Insecta); IMP.
DR GO; GO:0005072; P.ecdysone-mediated induction of salivary gla. .; NAS.
DR GO; GO:0007553; P.regulation of ecdysteroid metabolism; IMP.
DR InterPro; IPR001628; Hrm_rcpt_DNA_bd.
DR InterPro; IPR000536; Hrm_rcpt_lig_bd.
DR InterPro; IPR001723; Stdrim_receptor.
DR InterPro; IPR000324; Vld_receptor.
DR Pfam; PF00104; Hormone_recep; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRDHOMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR PRINTS; PR00350; VITAMINDR.
DR ProDom; PD000035; Znf_Cateteroid; 1.
DR SMART; SM00430; HOU1; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_REC_DBD_1; 1.
DR PROSITE; PS51030; NUCLEAR_REC_DBD_2; 1.
DR KMW Alternative splicing: Developmental protein; DNA-binding;
DR KMW Metal-binding; Nuclear protein; Receptor; Transcription;
DR KMW Transcription regulation; Zinc; Zinc-finger.
DR KMW Transcription regulation; Zinc; Zinc-finger.
DR FT DNA_BIND 455 531 NR C4-type 1.
DR FT ZN_FING 458 478 NR C4-type 2.
DR FT REGION 589 797 Ligand-binding (Potential).
DR FT COMPBIAS 6 10 Poly-Ala.
DR FT COMPBIAS 180 395 Gln-rich.
DR FT COMPBIAS 265 298 Pro-rich.
DR FT COMPBIAS 351 445 Ser-rich.
DR FT COMPBIAS 934 969 Gln-rich.
DR FT COMPBIAS 1113 1316 Ser-rich.
DR FT COMPFLCT 20 20 V -> F (in Ref. 1).
DR FT COMPFLCT 183 183 Q -> QQ (in Ref. 1).
DR FT COMPFLCT 337 342 Missing (in Ref. 1).
DR FT COMPFLCT 1142 1142 S -> V (in Ref. 1).
DR FT COMPFLCT 1216 1216 V -> L (in Ref. 1).
DR FT COMPFLCT 1231 1231 Missing (in Ref. 1).
DR FT COMPFLCT 1302 1302 N -> I (in Ref. 1).
DR SQ SEQUENCE 1412 AA; 151292 MW; 6F4C21B075443F8C CRC64;

Query Match 72.0%; Score 18; DB 1; Length 1412;
Best Local Similarity 22.2%; Pred. No. 66;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXX 9
Db 1283 CXXXXXXX 1291

RESULT 25
ID Q9W451 DROME PRELIMINARY; PRT; 1793 AA.
AC Q9W451;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE CG4790-PA.
GN Name=fe(1)M3; ORFNames=CG4790;
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
OX NCBI_TaxId=7227;
OX [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Gelinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champe M., Pfeiffer B.D.,

Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Agayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borikova K.C., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Fabros B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodok A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
 Idali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasio P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of *Drosophila melanogaster*."; *Science* 287:2185-2195(2000).
 [2]
 NUCLEOTIDE SEQUENCE.
 MEDLINE=22426065; PubMed=12537568;
 Celinker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
 Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 George R.A., Hoskins R.A., Laverly T., Murthy D.M., Nelson C.R.,
 Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 Weintraub G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
 melanogaster euchromatic genome sequence."; *Genome Biol.* 3:RESEARCH0079-RESEARCH0079(2002).
 [3]
 NUCLEOTIDE SEQUENCE.
 MEDLINE=22426070; PubMed=12537573;
 Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,
 Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 Ashburner M., Celinker S.E.;
 "The transposable elements of the *Drosophila melanogaster* euchromatin:
 a genomic perspective."; *Genome Biol.* 3:RESEARCH0084.1-RESEARCH0084.2(2002).
 [4]
 NUCLEOTIDE SEQUENCE.
 MEDLINE=22426069; PubMed=12537572;
 Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 Bettencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,
 Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
 Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 Lewis S.E.;
 "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 systematic review."; *Genome Biol.* 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 [5]
 NUCLEOTIDE SEQUENCE.
 Berkeley *Drosophila* Genome Project;
 Kaminker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
 Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
 Yu C., Rubin G.;

RT "Drosophila melanogaster release 4 sequence."; *GenBank/DBJ databases*.
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 [6]
 NUCLEOTIDE SEQUENCE.
 RP
 Submitter: (MAR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003436; AAF46107.1; -; Genomic DNA.
 DR Ensembl; CG4790; *Drosophila melanogaster*.
 DR FlyBase; FBgn005390; CG4790.
 DR FlyBase; FBgn005390; fs(1)M3.
 SQ SEQUENCE 1793 AA; 202233 MW; DDC38A764E3001CD CRC64;
 Query Match 72.0%; Score 18; DB 2; Length 1793;
 Best Local Similarity 22.2%; Pred. No. 73;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 CXXXXXXC 9
 DB 291 CTSSASASC 299
 RESULT 26
 Q8MP07_DROME
 ID Q8MP07_DROME PRELIMINARY; PRT; 1836 AA.
 AC Q8MP07;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Poliole precursor.
 GN Name=fs(1)M3; Synonyms=fs(1)ph;
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 OC NCBI_Taxid=7227;
 RX MEDLINE=21956567; PubMed=11959840; DOI=10.1101/gad.223902;
 RA Jimenez G., Gonzalez-Reyes A., Casanova J.;
 RT "Cell surface proteins Narsit and Poliole stabilize the Torso-like
 extracellular determinant in *Drosophila* oogenesis."; *Cell*
 RL Genes Dev. 16:913-918(2002).
 DR EMBL; AJ487522; CAD31790.1; -; mRNA.
 DR Ensembl; CG4790; *Drosophila melanogaster*.
 DR FlyBase; FBgn005390; CG4790.
 DR FlyBase; FBgn005390; fs(1)M3.
 KW Signal.
 FT SIGNAL 1 25 Potential.
 FT CHAIN 26 1836 Poliole.
 SQ SEQUENCE 1836 AA; 207449 MW; 4E0FC64DCB52EF CRC64;
 Query Match 72.0%; Score 18; DB 2; Length 1836;
 Best Local Similarity 22.2%; Pred. No. 74;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 CXXXXXXC 9
 DB 291 CTSSASASC 299
 RESULT 27
 Q4Q515_LEIMA
 ID Q4Q515_LEIMA PRELIMINARY; PRT; 1894 AA.
 AC Q4Q515;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Calpain-like cysteine peptidase, putative (Cysteine peptidase, clan
 CA, family C2, putative).
 GN ORFNames=LmIF32.0970;
 GN LmIF32.0970;
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OC NCBI_Taxid=5664;

```

RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Fiedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrett B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; C7005269; CAJ08617.1; -; Genomic DNA.
SQ SEQUENCE 1894 AA; 19683 MW; 3F71DE2F53BF95BB CRC64;

Query Match 72.0%; Score 18; DB 2; Length 1894;
Best Local Similarity 22.2%; Pred. No. 75;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 9
DB 1785 CAAAAAACC 1793

RESULT 28
ID Q4S488_TETNG PRELIMINARY; PRT; 2087 AA.
AC Q4S488;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Chromosome undetermined SCAP14743, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00024299001;
OS Tetradodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetradodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Michaud S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castell V., Katinka M., Vacherie B.,
RA Biemont C., Skallil Z., Cattolico L., Poulin J., De Bernardis V.,
RA Crnaud C., Duprat S., Broctier P., Couranceau J.P., Gouzy J.,
RA Paria G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetradodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01014743; CAG04544.1; -; Genomic DNA.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_1like.
DR InterPro; IPR000173; GAP_dhhdgenase.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR012680; Laminin_G_2.
DR Pfam; PF000028; Cadherin_8.
DR Pfam; PF01049; Cadherin_C1.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF02210; Laminin_G_2; 2.

```

```

DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 9.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00282; LamG; 2.
DR PROSITE; PS00232; CADHERIN_1; 6.
DR PROSITE; PS50268; CADHERIN_2; 9.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS00071; GAPDH; UNKNOWN_1.
DR PROSITE; PS50025; LamG_DOMAIN; 2.
KW Calcium; EGF-like domain; Transmembrane.
FT NON TER 1
SQ SEQUENCE 2087 AA; 22663 MW; 2C5D0554ECDD7A1 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 2087;
Best Local Similarity 22.2%; Pred. No. 78;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 9
DB 442 CSSSSAASC 450

RESULT 29
ID Q9L981_VIRCH PRELIMINARY; PRT; 50 AA.
AC Q9L981;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF31.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=569b;
RX MEDLINE=20476433; PubMed=11021935;
RA Clark C.A., Putins L., Kaewrakon P., Focareta T., Manning P.A.;
RL "The Vibrio cholerae O1 chromosomal integrin.";
RT Microbiology 146:2605-2612(2000).
DR EMBL; AF179596; AAF71194.1; -; Genomic DNA.
SQ SEQUENCE 50 AA; 5564 MW; 1418FC4DD3634E57 CRC64;

Query Match 68.0%; Score 17; DB 2; Length 50;
Best Local Similarity 22.2%; Pred. No. 57;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 9
DB 28 CASLSSSC 36

RESULT 30
ID Q4RA29_TETNG PRELIMINARY; PRT; 71 AA.
AC Q4RA29;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAP24619, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00037393001;
OS Tetradodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetradodon.
OX NCBI_TaxID=99883;
RN [1]

```

```

RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Mauceli E., Bouneau D., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Segrens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castell V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Bottier P., Coutanceau J.P., Gouzy J.,
RA Cruaud C., Duprat G., Brotier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chappie C., McKernan K.J., McSwan P., Bosak S.,
RA Kellis M., Volt J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN (2)
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01024619; CAG14754.1; -; Genomic_DNA.
FT NON TER 1 1
SQ SEQUENCE 71 AA; 8036 MW; 736AF2E349B9AC50 CRC64;

QY 1 CXXXXXXXC 9
Db 27 CSTRSPATC 35

RESULT 31
Q8VD96 MESAU PRELIMINARY; PRT; 96 AA.
ID Q8VD96;
AC Q8VD96;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE 10 kDa protein precursor.
GN Name=CC10;
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Cricetinae; Mesocricetus.
OX NCBI_TaxID=10036;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC Gutierrez-Sagal R., Nieto A.;
RT "Cloning and sequencing of a cDNA encoding hamster uteroglobin/clara
RT cell 10 kDa protein.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; L37041; AAL31349.1; -; Genomic_DNA.
DR HSSP; P17559; 1UTR.
DR SMR; Q8VD96; 20-96.
DR GO; GO:0005496; F:steroid binding; IEA.
DR InterPro; IPR006039; Utergl.
DR InterPro; IPR003628; Uteroglbn_sub.
DR InterPro; IPR000329; Uteroglbn_subf.
DR InterPro; IPR006038; Uteroglbn_subf.
DR PANTHER; PTHR10136; Uteroglbn_sub; 1.
DR Pfam; PF01039; Uteroglbn_sub; 1.
DR PRINTS; PR00486; UTEROGLBIN.
DR Prodom; PD012475; Uteroglbn_sub; 1.
DR SMART; SM00096; UTR; 1.
DR PROSITE; PS00404; UTEROGLBIN_2; 1.
KW Signal.
FT SIGNAL 1 18 Potential.

SQ SEQUENCE 96 AA; 10509 MW; 5EB9CBDD46143389 CRC64;
QY 1 CXXXXXXXC 9
Db 16 CSSASDTC 24

RESULT 32
Q4P4H6 USTWA PRELIMINARY; PRT; 108 AA.
ID Q4P4H6 USTWA PRELIMINARY;
AC Q4P4H6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Hypothetical protein.
GN ORFNames=UM04987.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Alt-Zehra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Amburster J., Bachmansang P., Baldwin J., Barry A.,
RA Bayul T., Blitshstein B., Bloom T., Biye J., Boguslavsky L.,
RA Borowsky M., Boukhalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Chestnang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawes T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Dufey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Geahrn G., Gnerre S.,
RA Gaitke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysagalis M., Karlsson E.,
RA Kells C., Kieu A., Kismet P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-Toh K., Liu X., Lokysang T., Lokysang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., MacDonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Marr K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., McGhee T., Melidrim J., Meneus L.,
RA Mesirov J., Mihaliev A., Mihova T., Mikkelson T., Mlenga V., Moru K.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pigani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Ratta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Ruttman M., Schupbach R., Seaman C., Settippalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Strange-Thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchniga P.,
RA Tenzing P., Teisfay S., Theodore J., Thoutang Y., Topham K.,
RA Towey S., Tsamila T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vanson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wengli T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis."
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACP01000184; EAK85760.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 108 AA; 11407 MW; E641B76D8BC16E7A CRC64;

```

Query Match 68.0%; Score 17; DB 2; Length 108;
 Best Local Similarity 22.2%; Pred. No. 78;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
 DB 14 CTSANATC 22

RESULT 33

Q9ZSP6_ARATH PRELIMINARY; PRT; 108 AA.

AC Q9ZSP6;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Putative cell wall-plasma membrane disconnecting CLCT protein.
 GN Name=ARL1B;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids 1; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;

RP NUCLEOTIDE SEQUENCE.
 RA Neutboom L.W., van der Zaal E.J.;
 RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.
 RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
 RA Nakajima M., Enju A., Kamiya A., Nannaka M., Carninci P., Kawai J.,
 RA Hayashizaki Y., Shinozaki K.;
 RL Submitted (Nov-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; APO98631; AAD12259.1; -; Genomic DNA.
 DR EMBL; AK118720; BAC43314.1; -; mRNA.
 DR HSSP; P24337; IHYP.
 DR InterPro; IPR003612; AAI.
 DR Pfam; PF00234; Tryp_alpha_amy1; 1.
 DR SMART; SM00499; AAI; 1.1194 MW; 8CD9F6CAE2B577D2 CRC64;
 SQ SEQUENCE 108 AA; 11194 MW; 8CD9F6CAE2B577D2 CRC64;

Query Match 68.0%; Score 17; DB 2; Length 108;
 Best Local Similarity 22.2%; Pred. No. 78;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
 DB 20 CTSATTGTC 28

RESULT 34

Q72D14_DESVH PRELIMINARY; PRT; 115 AA.

AC Q72D14;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedlocusNames=DVU1117;
 OS Desulfotribrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB 8303).
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfotribionales;
 OC Desulfotribriaceae; Desulfotribrio.
 OC NCBI_TaxID=882;

RP NUCLEOTIDE SEQUENCE.
 RA PubMed=15077118; DOI=10.1038/nbt959;
 RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
 RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
 RA Daugherty S.C., Deboy R.T., Dodson R.J., Durkin A.S., Madupu R.,
 RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
 RA Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D.,

RA Dimetrov G., Hance M., Tran K., Khouri H.M., Gill J., Uterback T.R.,
 RA Feldlyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
 RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
 RT Desulfotribrio vulgaris Hildenborough.";

RL Nat. Biotechnol. 22:554-559(2004).
 DR EMBL; AE017313; AAS95597.1; -; Genomic DNA.
 DR TIGR; DVU1117; -;
 KM Complete proteome; Hypothetical protein.
 SQ SEQUENCE 115 AA; 11838 MW; 8E8A4DDF3EB96E64 CRC64;

Query Match 68.0%; Score 17; DB 2; Length 115;
 Best Local Similarity 22.2%; Pred. No. 80;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
 DB 70 CDTATASC 78

RESULT 35

O6VZR1_CNPV PRELIMINARY; PRT; 117 AA.

AC O6VZR1;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE CNP0086 TNFR-like protein.
 GN Name=CNP0086;

OS Canaripox virus (CNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Avipoxvirus.
 OC NCBI_TaxID=44088;

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC VR-111;
 RA PubMed=14671117; DOI=10.1128/JVI.78.1.353-366.2004;
 RA Tulman E.R., Alonso C.L., Lu Z., Zsak J., Kutish G.F., Rock D.L.;
 RT "The genome of canaripox virus.";
 RL J. Virol. 78:353-366(2004).
 DR EMBL; AY318871; AAR83432.1; -; Genomic DNA.
 DR GO; GO:0004872; Fireceptor activity; IEA.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR001368; TNFR_c6.
 DR Pfam; PF00020; TNFR_c6; 1.
 DR SMART; SM00208; TNFR; 2.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
 DR PROSITE; PS50050; TNFR_NGFR_2; 1.
 SQ SEQUENCE 117 AA; 13113 MW; 06E99CA8A41135A CRC64;

Query Match 68.0%; Score 17; DB 2; Length 117;
 Best Local Similarity 22.2%; Pred. No. 80;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
 DB 50 CTSSTTTC 58

RESULT 36

Q4TWL2_ASF PRELIMINARY; PRT; 119 AA.

AC Q4TWL2;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE 9RL protein (Fragment).
 GN Name=B602L;
 OS African swine fever virus (ASFV).
 OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.

OC NCBI_TaxID=10497;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

```

RC STRAIN=Portalegre 90;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
  fever virus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028332; AAY46118.1; -; Genomic_DNA.
FT NON_TER 1 1
FT SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 22.2%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
DB 33 CASTGASTC 41

RESULT 37
Q4TWL3 ASF PRELIMINARY; PRT; 119 AA.
ID Q4TWL3;
AC Q4TWL3;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
ON NCBI_TaxID=10497;
RX [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Lisbon 60;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
  fever virus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028331; AAY46117.1; -; Genomic_DNA.
FT NON_TER 1 1
FT SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 22.2%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
DB 33 CASTGASTC 41

RESULT 38
Q4TWL5 ASF PRELIMINARY; PRT; 119 AA.
ID Q4TWL5;
AC Q4TWL5;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
ON NCBI_TaxID=10497;
RX [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Coimbra 87;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
  fever virus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028329; AAY46115.1; -; Genomic_DNA.
FT NON_TER 1 1
FT SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

```

```

FT NON_TER 119 119
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 22.2%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
DB 33 CASTGASTC 41

RESULT 39
Q4TWL6 ASF PRELIMINARY; PRT; 119 AA.
ID Q4TWL6;
AC Q4TWL6;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
ON NCBI_TaxID=10497;
RX [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Mafr 86;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
  fever virus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028328; AAY46114.1; -; Genomic_DNA.
FT NON_TER 1 1
FT SEQUENCE 119 AA; 11863 MW; 3A062A08CBA8D31F CRC64;

Query Match
Best Local Similarity 22.2%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
DB 33 CASTGASTC 41

RESULT 40
Q4TWL7 ASF PRELIMINARY; PRT; 119 AA.
ID Q4TWL7;
AC Q4TWL7;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
ON NCBI_TaxID=10497;
RX [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Barrancos 93;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
  fever virus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028327; AAY46113.1; -; Genomic_DNA.
FT NON_TER 1 1
FT SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 22.2%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 1 CXXXXXXC 9
DB 33 CASTGASTC 41

RESULT 41

Q4TWL8 ASF PRELIMINARY; PRT; 119 AA.
ID Q4TWL8 ASF PRELIMINARY; PRT; 119 AA.
AC Q4TWL8;
DT 13-SEP-2005 (TRENBLREL. 31, Created)
DT 13-SEP-2005 (TRENBLREL. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Almodovar 99;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine fever virus."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028326; AAY46112.1; -; Genomic_DNA.
FT NON TER 1
FT NON TER 119
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match 68.0%; Score 17; DB 2; Length 119;
Best Local Similarity 22.2%; Pred. No. 81;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 33 CASTGASTC 41

RESULT 42

Q4TWL9 ASF PRELIMINARY; PRT; 119 AA.
ID Q4TWL9 ASF PRELIMINARY; PRT; 119 AA.
AC Q4TWL9;
DT 13-SEP-2005 (TRENBLREL. 31, Created)
DT 13-SEP-2005 (TRENBLREL. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Almodovar 99/NE1;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine fever virus."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028325; AAY46111.1; -; Genomic_DNA.
FT NON TER 1
FT NON TER 119
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match 68.0%; Score 17; DB 2; Length 119;
Best Local Similarity 22.2%; Pred. No. 81;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 33 CASTGASTC 41

RESULT 43
Q4TWM0 ASF

ID Q4TWM0 ASF PRELIMINARY; PRT; 119 AA.

AC Q4TWM0;
DT 13-SEP-2005 (TRENBLREL. 31, Created)
DT 13-SEP-2005 (TRENBLREL. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Almodovar 99/E2;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine fever virus."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028324; AAY46110.1; -; Genomic_DNA.
FT NON TER 1
FT NON TER 119
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match 68.0%; Score 17; DB 2; Length 119;
Best Local Similarity 22.2%; Pred. No. 81;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 33 CASTGASTC 41

RESULT 44

O5YPN5 NOCFA PRELIMINARY; PRT; 125 AA.
ID O5YPN5 NOCFA PRELIMINARY; PRT; 125 AA.
AC O5YPN5;
DT 25-OCT-2004 (TRENBLREL. 28, Created)
DT 25-OCT-2004 (TRENBLREL. 28, Last sequence update)
DE Hypothetical protein.
GN Ordered locus names=nf50040;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K., Shiba T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IFM 10152."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR EMBL; AP006618; BAD59856.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 125 AA; 12749 MW; 1FD69A1CADF924EB CRC64;

Query Match 68.0%; Score 17; DB 2; Length 125;
Best Local Similarity 22.2%; Pred. No. 83;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
DB 96 CRSSSATTC 104

RESULT 45

Q4T2U4 TETNG PRELIMINARY; PRT; 133 AA.
ID Q4T2U4 TETNG PRELIMINARY; PRT; 133 AA.
AC Q4T2U4;
DT 13-SEP-2005 (TRENBLREL. 31, Created)
DT 13-SEP-2005 (TRENBLREL. 31, Last sequence update)
DE Chromosome 12 SCAP10174, whole genome shotgun sequence.

DE (Fragment).
 GN ORNames=GSTENG00008171001;
 OS Tetraodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 NC NCBITaxID=99883;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
 RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicard S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Dasilva C., Sallanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
 RA Biemont C., Skalli Z., Catolico L., Poulin J., De Berardinis V.,
 RA Cruaud C., Duprat S., Brothier P., Coutanceau J.P., Gouzy J.,
 RA Parra G., Lardier G., Chappie C., McKernan K.J., McEwan P., Bosak S.,
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Laudet V., Schacherer V., Quetier F., Saurin W., Scarpelli C.,
 RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.,
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 RT the early vertebrate proto-karyotype."
 RL Nature 431:946-957(2004).

RP NUCLEOTIDE SEQUENCE.
 RG Genoscope; Whitehead Institute Centre for Genome Research;
 RC Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL:CAAF010174; CAF92788.1; -; Genomic_DNA.
 FT NON_Coding
 SQ SEQUENCE 133 AA; 14711 MW; 528890284B8084F4 CRC64;

Query Match 68.0%; Score 17; DB 2; Length 133;
 Best Local Similarity 22.2%; Pred. No. 85;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
 DB 122 CQATSSPTC 130

RESULT 46
 Q6ZRF5 HUMAN PRELIMINARY; PRT; 134 AA.
 ID Q6ZRF5;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein FLJ46388.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;
 OC Homo.
 NC NCBITaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TSSUE=Thymus;
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
 RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
 RA Isono Y., Kawai-Hito Y., Saito K., Nishikawa T., Kimura K.,
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Yagatauma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
 RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahara K.,
 RA Masuno Y., Nagai K., Isogai T.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL:AKI28553; BAC87355.1; -; mRNA.
 SQ SEQUENCE 134 AA; 14481 MW; 03BA7E349C87BC6A CRC64;

Query Match 68.0%; Score 17; DB 2; Length 134;
 Best Local Similarity 22.2%; Pred. No. 85;
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9
 DB 27 CSASAAQC 35

RESULT 47
 Q9CWC3 MOUSE PRELIMINARY; PRT; 147 AA.
 ID Q9CWC3;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Mus musculus ES cells cDNA, RIKEN full-length enriched library,
 DE clone:2410101B22 product:LENG5 PROTEIN homolog.
 GN Name=Lang5;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 NC NCBITaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RT Carninci P., Hayashizaki Y.,
 RT "High-efficiency full-length cDNA cloning."
 RT Meth. Enzymol. 303:19-44(1999).
 RL [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=21085650; PubMed=11217851; DOI=10.1038/35055500;
 RT Kawai T., Shingawa A., Shidara K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsumoto H.A., Ashburner M., Batilov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glessl C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nishida T., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staudt L.F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bulc C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mondaerte P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J;
 RA The PANTOM Consortium,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/97.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes."
 RL Genome Res. 10:1617-1630(2000).
 RN [5]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CS7BL/6J;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Tashiro T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa K., Izawa M., Ohara E., Matshiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Ozaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CS7BL/6J;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arai K., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numata R., Ono M.,
RA Ozaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Takagawa A., Takahashi F., Tanaka T.,
RA Teijima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK010759; BAB27164.1; -, mRNA.
DR Ensembl; ENSMUSG00000035585; Mus musculus.
DR MGI; MGI:1913328; Leng5.
SO SEQUENCE 147 AA; 15856 MW; 988434686F9EA74 CRC64;
Query Match 68.0%; Score 17; DB 2; Length 147;
Best Local Similarity 22.2%; Pred. No. 88;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1 CXXXXXXC 9
Db 19 CATASTCTC 27
RESULT 48
Q6F340 ORYSA PRELIMINARY; PRT; 155 AA.
ID Q6F340 ORYSA PRELIMINARY; PRT; 155 AA.
AC Q6F340;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein OSUNBA0088106.14.
GN Name=OSUNBA0088106.14;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=39947;
OX [1]
RN NUCLEOTIDE SEQUENCE.
RP Chow T.-T., Hsiang Y.-I. C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Lue H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC129718; AAT69648.1; -, Genomic DNA.
DR Gramene; Q6F340; -;
KM Hypothetical protein.
SQ SEQUENCE 155 AA; 16979 MW; 0EADP4C906C39295 CRC64;
Query Match 68.0%; Score 17; DB 2; Length 155;
Best Local Similarity 22.2%; Pred. No. 90;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 9
Db 146 CRASASTC 154
RESULT 49
Q4TWL4 ASF PRELIMINARY; PRT; 145 AA.
ID Q4TWL4 ASF PRELIMINARY; PRT; 145 AA.
AC Q4TWL4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfarivirus.
OC NCBI_TaxID=10497;
OX [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NH/P68;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
RT fever virus."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028330; AAY46116.1; -, Genomic DNA.
FT NON_TER 1 1
FT NON_TER 155 155
SO SEQUENCE 155 AA; 15125 MW; 31AD6598783CE06 CRC64;
Query Match 68.0%; Score 17; DB 2; Length 155;
Best Local Similarity 22.2%; Pred. No. 90;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1 CXXXXXXC 9
Db 69 CASTGASTC 77
RESULT 50
CI025 HUMAN STANDARD; PRT; 157 AA.
ID CI025 HUMAN STANDARD; PRT; 157 AA.
AC Q81W50; Q5T591; Q8T2A2;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Protein C9orf25.
GN Name=C9orf25;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Eumarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OC NCBI_TaxID=9606;
OX [1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP PubMed=15164053; DOI=10.1038/nature02465;
RX PubMed=15164053; DOI=10.1038/nature02465;
RA Humphray S.J., Oliver K., Hunt A.R., Plumb R.W., Loveland J.B.,
RA Howe K.L., Andrews T.D., Searle S., Hunt S.E., Scott C.E., Jones M.C.,
RA Ainscough R., Almeida J.P., Ambrose K.D., Ashwell R.I.S.,
RA Barber D.J., Babbage S., Bagunley C.L., Bailey J., Banerjee R.,
RA Babbage A.K., Babbage K.F., Bates K., Beasley H., Beasley O., Bird C.P.,
RA Bray-Allen S., Brown A.J., Brown J.Y., Burford D., Burrill W.,
RA Burton J., Carder C., Carter N.P., Chapman J.C., Chen Y., Clarke G.,
RA Clark S.Y., Clegg C.M., Clegg S., Collier R.E., Corby N., Croxier M.,
RA Cummings A.T., Davies J., Dhami P., Dunn M., Dutta I., Dyer L.W.,
RA Earlrow M.E., Faulkner L., Fleming C.J., Frankish A.,
RA Frankland J.A., French L., Fricker D.G., Garner P., Garnett J.,
RA Ghori J., Gilbert J.G.R., Glison C., Grafham D.V., Griddle S.,
RA Griffiths C., Griffiths-Jones S., Grocock R., Guy J., Hall R.E.,
RA Hammond S., Harley J.L., Harrison B.S.I., Hart E.A., Heath P.D.,
RA Henderson C.D., Hopkins B.L., Howard P.J., Howden P.J., Huckle E.,
RA Johnson C., Johnson D., Joy A.A., Kay M., Keenan S., Kershaw J.K.,
RA Kimberley A.M., King A., Knights A., Laird G.K., Langford C.,

```

RA Lawlor S., Leongamornlert D.A., Leversha M., Lloyd C., Lloyd D.M.,
RA Lovell J., Martin S., Mashreghi-Mohammadi M., Matthews L., McLaren S.,
RA McLay K.E., McMurray A., Milne S., Nickerson T., Nisbett J.,
RA Nordiek G., Pearce A.V., Peck A.I., Porter K.M., Pandien R.,
RA Pelan S., Phillimore B., Povey S., Ramsey Y., Rand V., Scharfe M.,
RA Sehra H.K., Shownkeen R., Sims S.K., Skuce C.D., Smith M.,
RA Steward C.A., Swarbreck D., Symamore N., Tester J., Thorpe A.,
RA Tracey A., Tromans A., Thomas D.W., Wall M., Wallis J.M., West A.P.,
RA Whitehead S.L., Willey D.L., Williams S.A., Wilming L., Wray P.W.,
RA Young L., Aehret J.L., Coulson A., Blocker H., Durbin R.,
RA Sulston J.E., Hubbard T., Jackson M.J., Bentley D.R., Beck S.,
RA Rogers J., Dunham I.,
RT "DNA sequence and analysis of human chromosome 9."
RL Nature 429:369-374(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AL160270; CA113153.1; ALT INIT; Genomic DNA.
CC EMBL; AL156494; CA113153.1; JOINED; Genomic DNA.
CC EMBL; AL156494; CA115974.1; ALT INIT; Genomic DNA.
CC EMBL; AL160270; CA115974.1; JOINED; Genomic DNA.
CC EMBL; BC025348; AAH25348.1; -; mRNA.
CC EMBL; BC041009; AAH41009.1; -; mRNA.
CC EMBL; ENSG00000164970; Homo sapiens.
CC DR HGN; HGNC:19920; C9orf25.
CC DR HGN; HGNC:19920; C9orf25.
CC FT CONFLICT 106 106 Missing (in Ref. 2; AAH41009).
CC SQ SEQUENCE 157 AA; 17365 MW; 577AAFAA46B7DC74 CRC64;
CC -----
Query Match 68.0%; Score 17; DB 1; Length 157;
Best Local Similarity 22.2%; Pred. No. 90;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 CXXXXXXXXC 9
DB 146 COATSTRAC 154

```

Search completed: January 4, 2006, 16:09:53
 Job time : 59.7739 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 ; Search time 68.4348 Seconds
(without alignments)
64.204 Million cell updates/sec

Title: US-09-932-322-11

Perfect score: 26

Sequence: 1 CXXXXXXXXX 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

1: Geneseq_21.*
2: geneseqp1908.*
3: geneseqp1908.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	69.2	30	9	AD247471
2	18	69.2	89	8	ADP30533
3	18	69.2	89	8	ADP30531
4	18	69.2	109	8	AA093527
5	18	69.2	109	8	ADL31230
6	18	69.2	109	8	ADP30690
7	18	69.2	144	8	ADP31474
8	18	69.2	167	7	ABO69151
9	18	69.2	169	8	ADP312185
10	18	69.2	171	8	ADP30794
11	18	69.2	171	8	ADP30793
12	18	69.2	183	8	ADP30806
13	18	69.2	191	7	ABO76619
14	18	69.2	214	9	ADZ56213
15	18	69.2	228	8	ADP30921
16	18	69.2	252	8	ADP31485
17	18	69.2	258	8	ADP30479
18	18	69.2	264	8	ADP31412
19	18	69.2	270	8	ADP30500
20	18	69.2	270	8	ADP31321
21	18	69.2	294	8	ADP31473
22	18	69.2	297	8	ADP31192
23	18	69.2	328	4	AA025285
24	18	69.2	339	8	ADP30702

25	18	69.2	348	8	ADP31441	Adp31441 Human sec
26	18	69.2	356	4	ABG21039	Abg21039 Novel hum
27	18	69.2	357	8	ADP31267	Adp31267 Human sec
28	18	69.2	357	8	ADP30505	Adp30505 Human sec
29	18	69.2	360	8	ADP31439	Adp31439 Human sec
30	18	69.2	390	8	ADP31218	Adp31218 Human sec
31	18	69.2	421	8	ADP31159	Adp31159 Human sec
32	18	69.2	426	8	ADP31495	Adp31495 Human sec
33	18	69.2	438	7	ABO82561	AbO82561 Pseudomon
34	18	69.2	453	8	ADP31465	Adp31465 Human sec
35	18	69.2	471	8	ADP30854	Adp30854 Human sec
36	18	69.2	525	8	ADP31227	Adp31227 Human sec
37	18	69.2	549	8	ADP30855	Adp30855 Human sec
38	18	69.2	555	8	ADP31416	Adp31416 Human sec
39	18	69.2	555	8	ADP31417	Adp31417 Human sec
40	18	69.2	588	8	ADP31699	Adp31699 Human sec
41	18	69.2	591	9	ADZ56209	Adz56209 Human KIA
42	18	69.2	591	9	ADZ56223	Adz56223 Human KIA
43	18	69.2	604	8	ADP30865	Adp30865 Human sec
44	18	69.2	604	8	ADP30940	Adp30940 Human sec
45	18	69.2	604	8	ADP30941	Adp30941 Human sec
46	18	69.2	605	8	ADP30507	Adp30507 Human sec
47	18	69.2	617	8	ADP31657	Adp31657 Human sec
48	18	69.2	626	6	ABR41586	AbR41586 Human DIR
49	18	69.2	645	8	ADP30858	Adp30858 Human sec
50	18	69.2	653	3	AA9494907	AA9494907 Human sec
51	18	69.2	653	5	ADZ73254	Adz73254 Human KIA
52	18	69.2	653	5	ADW42108	Adw42108 Protein e
53	18	69.2	653	7	ADDB9077	TAT290. 1
54	18	69.2	653	8	ADZ510948	Adz510948 Human the
55	18	69.2	653	8	ADZ56204	Adz56204 Human KIA
56	18	69.2	653	9	ADZ56214	Adz56214 Human KIA
57	18	69.2	660	8	ADP31677	Adp31677 Human sec
58	18	69.2	665	8	ADP30571	Adp30571 Human sec
59	18	69.2	681	8	ADP31582	Adp31582 Human sec
60	18	69.2	690	8	ADP30780	Adp30780 Human sec
61	18	69.2	708	8	ADP31191	Adp31191 Human sec
62	18	69.2	711	8	ADP31440	Adp31440 Human sec
63	18	69.2	728	8	ADP30508	Adp30508 Human sec
64	18	69.2	739	8	ADP31196	Adp31196 Human sec
65	18	69.2	750	8	ADP30576	Adp30576 Human sec
66	18	69.2	757	8	ADP30925	Adp30925 Human sec
67	18	69.2	762	8	ADP30891	Adp30891 Human sec
68	18	69.2	771	8	ADP31469	Adp31469 Human sec
69	18	69.2	772	8	ADP30936	Adp30936 Human sec
70	18	69.2	774	8	ADP30506	Adp30506 Human sec
71	18	69.2	774	8	ADP30780	Adp30780 Human sec
72	18	69.2	774	8	ADP30781	Adp30781 Human sec
73	18	69.2	795	8	ADP31352	Adp31352 Human sec
74	18	69.2	804	8	ADP31291	Adp31291 Human sec
75	18	69.2	813	8	ADP31282	Adp31282 Human sec
76	18	69.2	813	8	ADP30649	Adp30649 Human sec
77	18	69.2	834	8	ADP30738	Adp30738 Human sec
78	18	69.2	852	8	ADP30664	Adp30664 Human sec
79	18	69.2	856	8	ADP30972	Adp30972 Human sec
80	18	69.2	865	8	ADP30566	Adp30566 Human sec
81	18	69.2	882	8	ADP31688	Adp31688 Human sec
82	18	69.2	888	8	ADP30971	Adp30971 Human sec
83	18	69.2	892	8	ADP31578	Adp31578 Human sec
84	18	69.2	930	8	ADP31444	Adp31444 Human sec
85	18	69.2	933	8	ADP31486	Adp31486 Human sec
86	18	69.2	947	8	ADP30937	Adp30937 Human sec
87	18	69.2	960	8	ADP31471	Adp31471 Human sec
88	18	69.2	960	8	ADP31470	Adp31470 Human sec
89	18	69.2	1017	6	AA026719	SR proteol
90	18	69.2	1033	8	ADP30984	Adp30984 Human sec
91	18	69.2	1065	8	ADP30661	Adp30661 Human sec
92	18	69.2	1065	8	ADP30973	Adp30973 Human sec
93	18	69.2	1065	8	ADP31482	Adp31482 Human sec
94	18	69.2	1090	8	ADQ10187	Adq10187 Human pol
95	18	69.2	1092	8	ADP31358	Adp31358 Human sec
96	18	69.2	1110	8	ADP31430	Adp31430 Human sec
97	18	69.2	1116	8	ADP31692	Adp31692 Human sec

98	18	69.2	1132	8	ADP30952	Adp30952 Human sec	171	18	69.2	3339	8	ADP31219	Adp31219 Human sec
99	18	69.2	1133	8	ADP30931	Adp30931 Human sec	172	18	69.2	3390	8	ADP31148	Adp31148 Human sec
100	18	69.2	1133	8	ADP30928	Adp30928 Human sec	173	18	69.2	3411	8	ADP30667	Adp30667 Human sec
101	18	69.2	1133	8	ADP30933	Adp30933 Human sec	174	18	69.2	3447	8	ADP31112	Adp31112 Human sec
102	18	69.2	1133	8	ADP30939	Adp30939 Human sec	175	18	69.2	3477	8	ADP30704	Adp30704 Human sec
103	18	69.2	1142	8	ADP30929	Adp30929 Human sec	176	18	69.2	4440	6	ABU88256	Abu88256 Novel hum
104	18	69.2	1165	8	ADP30879	Adp30879 Human sec	177	18	69.2	4440	6	ABU88256	Abu88256 Novel hum
105	18	69.2	1173	8	ADP31155	Adp31155 Human sec	178	18	69.2	4440	6	ABU96437	Abu96437 Novel hum
106	18	69.2	1173	8	ADP31034	Adp31034 Human sec	179	18	69.2	4440	6	ABU99046	Abu99046 Novel hum
107	18	69.2	1209	8	ADP30782	Adp30782 Human sec	180	18	69.2	4440	6	ABU98261	Abu98261 Novel hum
108	18	69.2	1222	8	ADP30501	Adp30501 Human sec	181	18	69.2	4440	6	ABU91967	Abu91967 Novel hum
109	18	69.2	1233	8	ADP30523	Adp30523 Human sec	182	18	69.2	4440	6	ABU85271	Abu85271 Novel hum
110	18	69.2	1239	8	ADP31297	Adp31297 Human sec	183	18	69.2	4440	6	ABU00410	Abu00410 Novel hum
111	18	69.2	1305	8	ADP31389	Adp31389 Human sec	184	18	69.2	4440	6	ABU88961	Abu88961 Novel hum
112	18	69.2	1358	8	ADP30995	Adp30995 Human sec	185	18	69.2	4440	6	ABU89517	Abu89517 Novel hum
113	18	69.2	1365	8	ADP31035	Adp31035 Human sec	186	18	69.2	4440	6	ABU95207	Abu95207 Novel hum
114	18	69.2	1371	8	ADP30876	Adp30876 Human sec	187	18	69.2	4440	6	ABU90755	Abu90755 Novel hum
115	18	69.2	1387	8	ADP30946	Adp30946 Human sec	188	18	69.2	4440	6	ABU93917	Abu93917 Novel hum
116	18	69.2	1431	8	ADP31609	Adp31609 Human sec	189	18	69.2	4440	6	ABU86191	Abu86191 Novel hum
117	18	69.2	1440	8	ADG34533	Adg34533 Glucocort	190	18	69.2	4440	6	ABU82046	Abu82046 Novel hum
118	18	69.2	1472	8	ADP31611	Adp31611 Human sec	191	18	69.2	4440	6	ABU07907	Abu07907 Novel hum
119	18	69.2	1480	8	ADP30557	Adp30557 Human sec	192	18	69.2	4440	6	ABU94227	Abu94227 Novel hum
120	18	69.2	1518	8	ADP31532	Adp31532 Human sec	193	18	69.2	4440	6	ABU00100	Abu00100 Novel hum
121	18	69.2	1550	8	ADP30567	Adp30567 Human sec	194	18	69.2	4440	6	ABU87111	Abu87111 Novel hum
122	18	69.2	1588	5	ABH09437	Abh09437 H. inflie	195	18	69.2	4440	6	ABU91352	Abu91352 Novel hum
123	18	69.2	1617	8	ADP30660	Adp30660 Human sec	196	18	69.2	4440	6	ABU90445	Abu90445 Novel hum
124	18	69.2	1771	8	ADP30674	Adp30674 Human sec	197	18	69.2	4440	6	ABU97036	Abu97036 Novel hum
125	18	69.2	1776	8	ADP31298	Adp31298 Human sec	198	18	69.2	4440	6	ABU05232	Abu05232 Novel hum
126	18	69.2	1776	9	ADY62715	Ady62715 Human alp	199	18	69.2	4440	6	ADP30585	Adp30585 Human sec
127	18	69.2	1782	8	ADP31391	Adp31391 Human sec	200	18	69.2	4752	8	ADP30651	Adp30651 Human sec
128	18	69.2	1782	8	ADP31270	Adp31270 Human sec	201	18	69.2	5304	8	ADP30705	Adp30705 Human sec
129	18	69.2	1794	8	ADP31176	Adp31176 Human sec	202	18	69.2	5304	8	ADP30705	Adp30705 Human sec
130	18	69.2	1878	8	ADP30740	Adp30740 Human sec	203	18	69.2	6729	8	ADP31600	Adp31600 Human sec
131	18	69.2	1894	6	ABU88253	Abu88253 Novel hum	204	18	69.2	6729	8	ADP30705	Adp30705 Human sec
132	18	69.2	1894	6	ABU90132	Abu90132 Novel hum	205	18	69.2	7285	6	ABJ38280	Abj38280 Human tca
133	18	69.2	1894	6	ABU96434	Abu96434 Novel hum	206	18	69.2	7339	6	AAOI6338	Aaoi6338 Human tca
134	18	69.2	1894	6	ABU99043	Abu99043 Novel hum	207	18	69.2	10944	8	ADP31310	Adp31310 Human sec
135	18	69.2	1894	6	ABU98258	Abu98258 Novel hum	208	18	69.2	11328	8	AAU00682	Aau00682 Human pro
136	18	69.2	1894	6	ABU91964	Abu91964 Novel hum	209	17	65.4	14	4	AAU020402	Aau020402 Human pep
137	18	69.2	1894	6	ABU85268	Abu85268 Novel hum	210	17	65.4	22	4	ABBA1096	Aba1096 Peptide #
138	18	69.2	1894	6	ABU88958	Abu88958 Novel hum	211	17	65.4	22	4	AAU34872	Aau34872 Peptide #
139	18	69.2	1894	6	ABU88958	Abu88958 Novel hum	212	17	65.4	22	4	ABBA2514	Aba2514 Peptide #
140	18	69.2	1894	6	ABU86188	Abu86188 Novel hum	213	17	65.4	22	4	ABBA2514	Aba2514 Peptide #
141	18	69.2	1894	6	ABU86188	Abu86188 Novel hum	214	17	65.4	22	4	AAU74756	Aau74756 Human bon
142	18	69.2	1894	6	ABU95204	Abu95204 Novel hum	215	17	65.4	22	4	AAU61952	Aau61952 Human bira
143	18	69.2	1894	6	ABU90752	Abu90752 Novel hum	216	17	65.4	22	4	ABG56539	Abg56539 Human bira
144	18	69.2	1894	6	ABU93914	Abu93914 Novel hum	217	17	65.4	22	4	AAU10165	Aau10165 Human bira
145	18	69.2	1894	6	ABU86188	Abu86188 Novel hum	218	17	65.4	22	4	AAU14688	Aau14688 Human bira
146	18	69.2	1894	6	ABU82043	Abu82043 Novel hum	219	17	65.4	22	4	AAU14688	Aau14688 Human bira
147	18	69.2	1894	6	ABU07904	Abu07904 Novel hum	220	17	65.4	22	4	AAU27105	Aau27105 Human bira
148	18	69.2	1894	6	ABU94224	Abu94224 Novel hum	221	17	65.4	22	4	ABBA28466	Aba28466 Peptide #
149	18	69.2	1894	6	ABU00097	Abu00097 Novel hum	222	17	65.4	22	4	ABBA28466	Aba28466 Peptide #
150	18	69.2	1894	6	ABU87108	Abu87108 Novel hum	223	17	65.4	22	4	ABBA28466	Aba28466 Peptide #
151	18	69.2	1894	6	ABU91349	Abu91349 Novel hum	224	17	65.4	22	4	ABBA28466	Aba28466 Peptide #
152	18	69.2	1894	6	ABU90442	Abu90442 Novel hum	225	17	65.4	22	4	AAU66822	Aau66822 Human bon
153	18	69.2	1894	6	ABU97033	Abu97033 Novel hum	226	17	65.4	22	4	AAU66822	Aau66822 Human bon
154	18	69.2	1894	6	ABU05229	Abu05229 Novel hum	227	17	65.4	22	4	ABG48490	Abg48490 Human bira
155	18	69.2	1917	6	ADAI5719	Adai5719 C. elegan	228	17	65.4	52	4	AAU02411	Aau02411 Peptide #
156	18	69.2	2088	8	ADP31578	Adp31578 Human sec	229	17	65.4	52	4	ABG36477	Abg36477 Human pep
157	18	69.2	2088	8	ADP31178	Adp31178 Human sec	230	17	65.4	52	4	ABG36477	Abg36477 Human pep
158	18	69.2	2187	8	ADP30882	Adp30882 Human sec	231	17	65.4	62	4	AAU28428	Aau28428 Peptide #
159	18	69.2	2272	8	ADP30669	Adp30669 Human sec	232	17	65.4	62	4	ABBA29741	Aba29741 Peptide #
160	18	69.2	2304	8	ADP31252	Adp31252 Human sec	233	17	65.4	62	4	ABBA29741	Aba29741 Peptide #
161	18	69.2	2307	8	ADP31394	Adp31394 Human sec	234	17	65.4	62	4	AAU68101	Aau68101 Human bon
162	18	69.2	2415	8	ADP31023	Adp31023 Human sec	235	17	65.4	62	4	AAU68101	Aau68101 Human bon
163	18	69.2	2469	8	ADP31031	Adp31031 Human sec	236	17	65.4	62	4	ABG49751	Abg49751 Human liv
164	18	69.2	2542	8	ADP31594	Adp31594 Human sec	237	17	65.4	62	4	AAU03661	Aau03661 Peptide #
165	18	69.2	2616	8	ADP31253	Adp31253 Human sec	238	17	65.4	62	4	ABG37626	Abg37626 Human pep
166	18	69.2	2616	8	ABBA4675	Aba4675 N. mening	239	17	65.4	68	4	AAU14223	Aau14223 Peptide #
167	18	69.2	2980	9	AEA36049	Aea36049 Walze Sta	240	17	65.4	68	4	AAU14223	Aau14223 Peptide #
168	18	69.2	3036	8	ADP31595	Adp31595 Human sec	241	17	65.4	68	4	AAU26633	Aau26633 Peptide #
169	18	69.2	3070	8	ADP30710	Adp30710 Human sec	242	17	65.4	68	4	ABBA27998	Aba27998 Human pep
170	18	69.2	3070	8	ADP30700	Adp30700 Human sec	243	17	65.4	68	4	ABBA27998	Aba27998 Human pep

390	17	65.4	618	8	ADP31554	Adp31554 Human sec	463	17	65.4	996	8	ADP31538	Adp31538 Human sec
391	17	65.4	621	8	ADP31147	Adp31147 Human sec	464	17	65.4	1017	4	AB59813	Adp59813 Tyld prot
392	17	65.4	624	8	ADP31324	Adp31324 Human sec	465	17	65.4	1017	8	ADP31268	Adp31268 Human sec
393	17	65.4	624	8	ADP31325	Adp31325 Human sec	466	17	65.4	1035	8	ADP31552	Adp31552 Human sec
394	17	65.4	629	8	ADP31295	Adp31295 Human sec	467	17	65.4	1041	8	ADP30768	Adp30768 Human sec
395	17	65.4	639	8	ADP30861	Adp30861 Human sec	468	17	65.4	1044	8	ADP31517	Adp31517 Human sec
396	17	65.4	639	9	AD210464	Ad210464 P. glingiv	469	17	65.4	1050	8	ADP31370	Adp31370 Human sec
397	17	65.4	642	8	ADP31161	Adp31161 Human sec	470	17	65.4	1053	8	ADP30886	Adp30886 Human sec
398	17	65.4	645	8	ADP31124	Adp31124 Human sec	471	17	65.4	1056	8	ADP31082	Adp31082 Human sec
399	17	65.4	645	8	ADP31125	Adp31125 Human sec	472	17	65.4	1059	8	ADP31042	Adp31042 Human sec
400	17	65.4	654	8	ADP31102	Adp31102 Human sec	473	17	65.4	1062	8	ADP31369	Adp31369 Human sec
401	17	65.4	660	8	ADP31606	Adp31606 Human sec	474	17	65.4	1066	8	ADP30563	Adp30563 Human sec
402	17	65.4	666	8	ADP30466	Adp30466 Human sec	475	17	65.4	1076	6	ABU81145	Abu81145 Human PRO
403	17	65.4	669	8	ADP31598	Adp31598 Human sec	476	17	65.4	1076	6	ABU66845	Abu66845 Human PRO
404	17	65.4	669	8	ADP31493	Adp31493 Human sec	477	17	65.4	1083	8	ADP30473	Adp30473 Human sec
405	17	65.4	670	8	ADP31001	Adp31001 Human sec	478	17	65.4	1086	8	ADP31175	Adp31175 Human sec
406	17	65.4	672	8	ADP30498	Adp30498 Human sec	479	17	65.4	1086	8	ADP31447	Adp31447 Human sec
407	17	65.4	681	8	ADP31053	Adp31053 Human sec	480	17	65.4	1086	8	ADP31629	Adp31629 Human sec
408	17	65.4	687	8	ADP31364	Adp31364 Human sec	481	17	65.4	1087	8	ADP30978	Adp30978 Human sec
409	17	65.4	705	8	ADP31623	Adp31623 Human sec	482	17	65.4	1104	8	ADP31156	Adp31156 Human sec
410	17	65.4	711	8	ADP31235	Adp31215 Human sec	483	17	65.4	1113	8	ADP31508	Adp31508 Human sec
411	17	65.4	711	8	ADP31535	Adp31535 Human sec	484	17	65.4	1116	8	ADP31203	Adp31203 Human sec
412	17	65.4	714	8	ADP31561	Adp31561 Human sec	485	17	65.4	1116	8	ADP31202	Adp31202 Human sec
413	17	65.4	717	8	ADP31410	Adp31410 Human sec	486	17	65.4	1128	8	ADP30483	Adp30483 Human sec
414	17	65.4	725	8	ADP31092	Adp31092 Human sec	487	17	65.4	1134	8	ADP31537	Adp31537 Human sec
415	17	65.4	732	8	ADP30650	Adp30650 Human sec	488	17	65.4	1141	8	ADP30677	Adp30677 Human sec
416	17	65.4	746	8	ADP30982	Adp30982 Human sec	489	17	65.4	1147	8	ADP30965	Adp30965 Human sec
417	17	65.4	750	8	ADP31131	Adp31131 Human sec	490	17	65.4	1155	8	ADP31455	Adp31455 Human sec
418	17	65.4	755	8	ADP31559	Adp31559 Human sec	491	17	65.4	1168	8	ADP31046	Adp31046 Human sec
419	17	65.4	756	8	ADP31039	Adp31039 Human sec	492	17	65.4	1189	8	ADP31043	Adp31043 Human sec
420	17	65.4	764	8	ADP31614	Adp31614 Human sec	493	17	65.4	1191	8	ADP30993	Adp30993 Human sec
421	17	65.4	765	8	ADP31149	Adp31149 Human sec	494	17	65.4	1192	8	ADP31180	Adp31180 Human sec
422	17	65.4	774	8	ADP30885	Adp30885 Human sec	495	17	65.4	1199	8	ADP31044	Adp31044 Human sec
423	17	65.4	774	8	ADP31373	Adp31373 Human sec	496	17	65.4	1200	8	ADP31603	Adp31603 Human sec
424	17	65.4	779	8	ADP30915	Adp30915 Human sec	497	17	65.4	1221	8	ADP31293	Adp31293 Human sec
425	17	65.4	779	8	ADP30897	Adp30897 Human sec	498	17	65.4	1224	8	ADP31426	Adp31426 Human sec
426	17	65.4	780	8	ADP31077	Adp31077 Human sec	499	17	65.4	1227	8	ADP31602	Adp31602 Human sec
427	17	65.4	783	8	ADP31398	Adp31398 Human sec	500	17	65.4	1227	8	ADP31210	Adp31210 Human sec
428	17	65.4	804	8	ADP31635	Adp31635 Human sec	501	17	65.4	1248	8	ADP31346	Adp31346 Human sec
429	17	65.4	806	8	ADP31110	Adp31110 Human sec	502	17	65.4	1252	8	ADP30678	Adp30678 Human sec
430	17	65.4	813	8	ADP30561	Adp30561 Human sec	503	17	65.4	1260	8	ADP31533	Adp31533 Human sec
431	17	65.4	816	8	ADP31249	Adp31249 Human sec	504	17	65.4	1269	8	ADP31382	Adp31382 Human sec
432	17	65.4	816	8	ADP31581	Adp31581 Human sec	505	17	65.4	1269	8	ADP31381	Adp31381 Human sec
433	17	65.4	823	4	ABG25826	Abg25826 Novel hum	506	17	65.4	1282	8	ADP31328	Adp31328 Human sec
434	17	65.4	831	8	ADN22420	Adn22420 Bacterial	507	17	65.4	1288	8	ADP31114	Adp31114 Human sec
435	17	65.4	831	8	ADP31333	Adp31333 Human sec	508	17	65.4	1289	8	ADP30675	Adp30675 Human sec
436	17	65.4	831	8	ADP31179	Adp31179 Human sec	509	17	65.4	1300	6	ABU88254	Abu88254 Novel hum
437	17	65.4	876	8	ADP31220	Adp31220 Human sec	510	17	65.4	1300	6	ABU90133	Abu90133 Novel hum
438	17	65.4	887	8	ADP30554	Adp30554 Human sec	511	17	65.4	1300	6	ABU96435	Abu96435 Novel hum
439	17	65.4	887	8	ADP30548	Adp30548 Human sec	512	17	65.4	1300	6	ABU99044	Abu99044 Novel hum
440	17	65.4	888	8	ADP31648	Adp31648 Human sec	513	17	65.4	1300	6	ABU98259	Abu98259 Novel hum
441	17	65.4	890	8	ADP31059	Adp31059 Human sec	514	17	65.4	1300	6	ABU91965	Abu91965 Novel hum
442	17	65.4	897	8	ADP30914	Adp30914 Human sec	515	17	65.4	1300	6	ABU85269	Abu85269 Novel hum
443	17	65.4	903	8	ADP31294	Adp31294 Human sec	516	17	65.4	1300	6	ABU00408	Abu00408 Novel hum
444	17	65.4	906	6	ADP31573	Adp31573 Human sec	517	17	65.4	1300	6	ABU88959	Abu88959 Novel hum
445	17	65.4	908	6	ABU08492	Abu08492 ALpha-hel	518	17	65.4	1300	6	ABO06455	Abu06455 Novel hum
446	17	65.4	912	8	ADP31507	Adp31507 Human sec	519	17	65.4	1300	6	ABU95515	Abu95515 Novel hum
447	17	65.4	918	8	ADP31459	Adp31459 Human sec	520	17	65.4	1300	6	ABU95205	Abu95205 Novel hum
448	17	65.4	921	8	ADP31522	Adp31522 Human sec	521	17	65.4	1300	6	ABU90753	Abu90753 Novel hum
449	17	65.4	925	5	AAO14246	AAo14246 Human pre	522	17	65.4	1300	6	ABU93915	Abu93915 Novel hum
450	17	65.4	933	8	ADP31510	Adp31510 Human sec	523	17	65.4	1300	6	ABU86189	Abu86189 Novel hum
451	17	65.4	936	8	ADP31597	Adp31597 Human sec	524	17	65.4	1300	6	ABU82044	Abu82044 Novel hum
452	17	65.4	939	8	ADP31086	Adp31086 Human sec	525	17	65.4	1300	6	ABU07905	Abu07905 Novel hum
453	17	65.4	939	8	ADP30726	Adp30726 Human sec	526	17	65.4	1300	6	ABU94225	Abu94225 Novel hum
454	17	65.4	942	8	ADN22417	Adn22417 Bacterial	527	17	65.4	1300	6	ABO00098	Abu000098 Novel hum
455	17	65.4	948	8	ADP30586	Adp30586 Human sec	528	17	65.4	1300	6	ABU87109	Abu87109 Novel hum
456	17	65.4	951	8	ADP31413	Adp31413 Human sec	529	17	65.4	1300	6	ABU91350	Abu91350 Novel hum
457	17	65.4	966	8	ADP30745	Adp30745 Human sec	530	17	65.4	1300	6	ABU90443	Abu90443 Novel hum
458	17	65.4	967	8	ADP30698	Adp30698 Human sec	531	17	65.4	1300	6	ABU97034	Abu97034 Novel hum
459	17	65.4	967	8	ADP30718	Adp30718 Human sec	532	17	65.4	1302	8	ABO05230	Abu005230 Novel hum
460	17	65.4	976	8	ADP30688	Adp30688 Human sec	533	17	65.4	1302	8	ADP31695	Adp31695 Human sec
461	17	65.4	981	8	ADP30547	Adp30547 Human sec	534	17	65.4	1303	8	ADP30751	Adp30751 Human sec
462	17	65.4	990	8	ADP31553	Adp31553 Human sec	535	17	65.4	1312	8	ADP30999	Adp30999 Human sec

536	17	65.4	1314	8	ADP31197	Human sec	609	17	65.4	1803	8	ADP30562	Adp30562 Human sec
537	17	65.4	1344	8	ADP31211	Human sec	610	17	65.4	1815	8	ADP31601	Adp31601 Human sec
538	17	65.4	1365	8	ADP31661	Human sec	611	17	65.4	1827	8	ADP31170	Adp31170 Human sec
539	17	65.4	1371	8	ADP31646	Human sec	612	17	65.4	1833	8	ADP30642	Adp30642 Human sec
540	17	65.4	1380	8	ADP31566	Human sec	613	17	65.4	1848	8	ADP31372	Adp31372 Human sec
541	17	65.4	1386	8	ADP31371	Human sec	614	17	65.4	1849	8	ADP30752	Adp30752 Human sec
542	17	65.4	1398	8	ADP31488	Human sec	615	17	65.4	1875	8	ADP31664	Adp31664 Human sec
543	17	65.4	1401	8	ADP30532	Human sec	616	17	65.4	1933	8	ADP30889	Adp30889 Human sec
544	17	65.4	1417	8	ADP31160	Human sec	617	17	65.4	1933	8	ADP30902	Adp30902 Human sec
545	17	65.4	1437	8	ADP31357	Human sec	618	17	65.4	1966	8	ADP31662	Adp31662 Human sec
546	17	65.4	1454	8	ADP31177	Human sec	619	17	65.4	1968	8	ADP30689	Adp30689 Human sec
547	17	65.4	1456	8	ADP30923	Human sec	620	17	65.4	2001	8	ADP31644	Adp31644 Human sec
548	17	65.4	1464	8	ADP31040	Human sec	621	17	65.4	2020	8	ADP31056	Adp31056 Human sec
549	17	65.4	1464	8	ADP31437	Human sec	622	17	65.4	2027	8	ADP31058	Adp31058 Human sec
550	17	65.4	1485	8	ADP31383	Human sec	623	17	65.4	2052	9	AEH87635	Adp30687 Human sec
551	17	65.4	1485	8	ADP31384	Human sec	624	17	65.4	2058	8	ADP31650	Adp31650 Human sec
552	17	65.4	1488	8	ADP31385	Human sec	625	17	65.4	2123	8	ADP30657	Adp30657 Human sec
553	17	65.4	1488	8	ADP31386	Human sec	626	17	65.4	2123	8	ADP30564	Adp30564 Human sec
554	17	65.4	1494	8	ADP31650	Human sec	627	17	65.4	2127	8	ADP31327	Adp31327 Human sec
555	17	65.4	1506	8	ADP30596	Human sec	628	17	65.4	2148	8	ADP30974	Adp30974 Human sec
556	17	65.4	1521	8	ADP30542	Human sec	629	17	65.4	2260	8	ADP30687	Adp30687 Human sec
557	17	65.4	1521	8	ADP30537	Human sec	630	17	65.4	2272	8	ADP31136	Adp31136 Human sec
558	17	65.4	1530	8	ADP31536	Human sec	631	17	65.4	2349	8	ADP30959	Adp30959 Human sec
559	17	65.4	1536	8	ADP31302	Human sec	632	17	65.4	2382	8	ADP31341	Adp31341 Human sec
560	17	65.4	1539	8	ADP31201	Human sec	633	17	65.4	2454	8	ADP30469	Adp30469 Human sec
561	17	65.4	1560	8	ADP31200	Human sec	634	17	65.4	2484	8	ADP66690	Adp66690 Human sec
562	17	65.4	1560	8	ADP31570	Human sec	635	17	65.4	2508	6	ADA15721	Adp66690 Human mis
563	17	65.4	1560	8	ADP31139	Human sec	636	17	65.4	2535	8	ADP31146	Adp31146 Human sec
564	17	65.4	1560	8	ADP31631	Human sec	637	17	65.4	2544	6	ADA15717	Adp31146 Human sec
565	17	65.4	1566	8	ADP31407	Human sec	638	17	65.4	2547	8	ADP31665	Adp31665 Human sec
566	17	65.4	1584	8	ADP31405	Human sec	639	17	65.4	2601	6	ADA15723	Adp31665 Human sec
567	17	65.4	1587	8	ADP30591	Human sec	640	17	65.4	2622	8	ADP31673	Adp31673 Human sec
568	17	65.4	1596	8	ADP31491	Human sec	641	17	65.4	2664	8	ADP31546	Adp31546 Human sec
569	17	65.4	1635	4	AAH59826	Protein #	642	17	65.4	2688	8	ADN11593	Adp31546 Human sec
570	17	65.4	1623	8	ADP30552	Human sec	643	17	65.4	2724	8	ADP31233	Adp31233 Human sec
571	17	65.4	1631	4	ABG22481	Novel hum	644	17	65.4	2790	8	ADP31232	Adp31232 Human sec
572	17	65.4	1647	8	ADP30670	Human sec	645	17	65.4	2808	8	ADP31258	Adp31258 Human sec
573	17	65.4	1647	8	ADP31052	Human sec	646	17	65.4	2827	8	ADP30681	Adp30681 Human sec
574	17	65.4	1652	6	ADA15715	C. elegans	647	17	65.4	2828	8	ADP30938	Adp30938 Human sec
575	17	65.4	1654	8	ADP30997	Human sec	648	17	65.4	2833	8	ADP31299	Adp31299 Human sec
576	17	65.4	1662	8	ADP31419	Human sec	649	17	65.4	2835	8	ADP30572	Adp30572 Human sec
577	17	65.4	1662	8	ADP31513	Human sec	650	17	65.4	3046	8	ADQ10419	Human pol
578	17	65.4	1665	8	ADP31187	Human sec	651	17	65.4	3144	8	ADP31544	Adp31544 Human sec
579	17	65.4	1700	3	AAH18144	Plasmodiu	652	17	65.4	3201	8	ADP31545	Adp31545 Human sec
580	17	65.4	1725	8	ADP30654	Human sec	653	17	65.4	3398	9	AEH87634	Human ino
581	17	65.4	1737	8	ADP31292	Human sec	654	17	65.4	3407	8	ADP31060	Adp31060 Human sec
582	17	65.4	1743	6	ABH88255	Novel hum	655	17	65.4	3407	8	ADP31062	Adp31062 Human sec
583	17	65.4	1743	6	ABH90134	Novel hum	656	17	65.4	3465	8	ADP31234	Adp31234 Human sec
584	17	65.4	1743	6	ABH96436	Novel hum	657	17	65.4	3579	8	ADP31098	Adp31098 Human sec
585	17	65.4	1743	6	ABH99045	Novel hum	658	17	65.4	3638	8	ADP30981	Adp30981 Human sec
586	17	65.4	1743	6	ABH98260	Novel hum	659	17	65.4	4683	8	ADP31260	Adp31260 Human sec
587	17	65.4	1743	6	ABH91966	Novel hum	660	17	65.4	4848	8	ADP31259	Adp31259 Human sec
588	17	65.4	1743	6	ABH85270	Novel hum	661	17	65.4	5397	8	ADP31068	Adp31068 Human sec
589	17	65.4	1743	6	ABH00409	Novel hum	662	17	65.4	5514	8	ADP31186	Adp31186 Human sec
590	17	65.4	1743	6	ABH88960	Novel hum	663	17	65.4	5514	8	ADP31591	Adp31591 Human sec
591	17	65.4	1743	6	ABH06456	Novel hum	664	17	65.4	5820	8	ADP31118	Adp31118 Human sec
592	17	65.4	1743	6	ABH95516	Novel hum	665	17	65.4	8973	8	ADP31119	Adp31119 Human sec
593	17	65.4	1743	6	ABH95206	Novel hum	666	17	65.4	8976	8	ADP31425	Adp31425 Human sec
594	17	65.4	1743	6	ABH90754	Novel hum	667	17	65.4	9195	8	ADP31494	Adp31494 Human sec
595	17	65.4	1743	6	ABH93916	Novel hum	668	17	65.4	10	2	AAH71548	Adp31494 Human sec
596	17	65.4	1743	6	ABH86190	Novel hum	669	17	65.4	10	2	AAH28918	Adp31494 Human sec
597	17	65.4	1743	6	ABH82045	Novel hum	670	17	65.4	10	2	AAH805791	Adp31494 Human sec
598	17	65.4	1743	6	ABH07906	Novel hum	671	17	65.4	22	9	AAH76485	Adp31494 Human sec
599	17	65.4	1743	6	ABH94226	Novel hum	672	17	65.4	22	5	AAH21859	Adp31494 Human sec
600	17	65.4	1743	6	ABH00099	Novel hum	673	17	65.4	25	5	AAH19915	Adp31494 Human sec
601	17	65.4	1743	6	ABH087110	Novel hum	674	17	65.4	25	7	AAH00366	Adp31494 Human sec
602	17	65.4	1743	6	ABH91351	Novel hum	675	17	65.4	25	8	ADG47681	Adp31494 Human sec
603	17	65.4	1743	6	ABH90444	Novel hum	676	17	65.4	34	5	AAH87022	Adp31494 Human sec
604	17	65.4	1743	6	ABH97035	Novel hum	677	17	65.4	41	4	AAH41607	Adp31494 Human sec
605	17	65.4	1743	6	ABH005231	Novel hum	678	17	65.4	41	4	AAH35400	Adp31494 Human sec
606	17	65.4	1749	8	ADP31408	Human sec	679	17	65.4	41	4	AAH25433	Adp31494 Human sec
607	17	65.4	1755	8	ADP31446	Human sec	680	17	65.4	41	4	AAH75287	Adp31494 Human sec
608	17	65.4	1789	8	ADP30962	Human sec	681	17	65.4	41	4	AAH62478	Adp30962 Human sec

682	16	61.5	41	4	ABG57048	Abg57048 Human liv	755	16	61.5	110	6	ABM65559	Abm65559 Propionib
683	16	61.5	41	5	ABG44927	Abg44927 Human pep	756	16	61.5	111	5	ADK35517	Adk35517 Novel hum
684	16	61.5	44	2	AAR98208	Aar98208 Nucleotid	757	16	61.5	112	7	ABO68317	AbO68317 Pseudomon
685	16	61.5	45	8	ADT92777	Adt92777 Mouse Cal	758	16	61.5	112	8	ADX88804	Adx88804 Plant ful
686	16	61.5	45	8	ADT92788	Adt92788 Mouse Cal	759	16	61.5	113	8	ABT03134	AbT03134 SARS coro
687	16	61.5	47	9	ADZ47821	Adz47821 Antno aci	760	16	61.5	113	9	AEB91569	Aeb91569 Microbial
688	16	61.5	48	3	AGS57161	AgS57161 Arabidops	761	16	61.5	114	5	ABP06472	Abp06472 Human ORF
689	16	61.5	48	3	AMG50935	Amg50935 Arabidops	762	16	61.5	115	4	AAH83845	Aah83845 Human imm
690	16	61.5	48	9	ADV76464	Adv76464 Human HER	763	16	61.5	116	8	ADP30955	Adp30955 Human sec
691	16	61.5	48	9	AEA34233	Aea34233 Optum pop	764	16	61.5	118	4	ABG22738	Abg22738 Novel hum
692	16	61.5	51	8	ADP29386	Adp29386 Human sec	765	16	61.5	121	9	ABM96843	Abm96843 M. xanthu
693	16	61.5	51	8	ADP29383	Adp29383 Human sec	766	16	61.5	122	4	ABBI17561	Abbi17561 Human ner
694	16	61.5	52	5	ADP31338	Adp31338 Human ORF	767	16	61.5	122	7	ABO83310	AbO83310 Pseudomon
695	16	61.5	58	4	AAU48534	Aau48534 Propionib	768	16	61.5	123	4	AAH90307	Aah90307 Human imm
696	16	61.5	58	6	ABM45053	Abm45053 Propionib	769	16	61.5	125	4	ABG04861	Abg04861 Novel hum
697	16	61.5	58	6	ABM45053	Abm45053 Propionib	770	16	61.5	133	6	AAU64082	Aau64082 Propionib
698	16	61.5	61	4	AAH19250	Aah19250 Peptide #	771	16	61.5	133	4	ABM60601	Abm60601 Human sec
699	16	61.5	61	4	ABM31941	Abm31941 Peptide #	772	16	61.5	133	8	ADP31069	Adp31069 Human sec
700	16	61.5	61	4	ABM31941	Abm31941 Peptide #	773	16	61.5	134	7	ABO75321	AbO75321 Pseudomon
701	16	61.5	61	4	ABM31941	Abm31941 Peptide #	774	16	61.5	136	7	ABO73061	AbO73061 Pseudomon
702	16	61.5	61	4	AAH71647	Aah71647 Human bon	775	16	61.5	137	4	AAH93766	Aah93766 Human pro
703	16	61.5	61	4	AAH59112	Aah59112 Human bra	776	16	61.5	137	7	ABO82722	AbO82722 Pseudomon
704	16	61.5	61	4	ABG53331	Abg53331 Human liv	777	16	61.5	139	7	ABO80767	AbO80767 Pseudomon
705	16	61.5	63	4	AAU59447	Aau59447 Human pep	778	16	61.5	139	7	ABO72721	AbO72721 Pseudomon
706	16	61.5	63	6	ABM55966	Abm55966 Propionib	779	16	61.5	139	7	ABO76095	AbO76095 Pseudomon
707	16	61.5	63	6	ABR63898	AbR63898 P. furiosus	780	16	61.5	140	7	ABO83386	AbO83386 Pseudomon
708	16	61.5	65	4	ABP35884	Abp35884 CTRB_GADH	781	16	61.5	140	7	ABG63884	AbG63884 Human pro
709	16	61.5	65	4	ABP35883	Abp35883 CTRB_GADH	782	16	61.5	141	4	ABO84154	AbO84154 Pseudomon
710	16	61.5	65	4	AAU41244	Aau41244 Propionib	783	16	61.5	146	7	ABO72664	AbO72664 Pseudomon
711	16	61.5	65	6	ABM37763	Abm37763 Propionib	784	16	61.5	146	7	ABG29376	AbG29376 Human SAR
712	16	61.5	66	8	AAH57014	Aah57014 Arabidops	785	16	61.5	147	4	ABO79762	AbO79762 Pseudomon
713	16	61.5	68	8	ADP31694	Adp31694 Human sec	786	16	61.5	150	2	AAH49809	Aah49809 Human g11
714	16	61.5	70	7	ABO73323	AbO73323 Pseudomon	787	16	61.5	150	3	AAH20620	Aah20620 PTM3 huma
715	16	61.5	71	4	AAH14601	Aah14601 Peptide #	788	16	61.5	150	7	ABO64648	AbO64648 Human sec
716	16	61.5	71	4	ABM33561	Abm33561 Peptide #	789	16	61.5	150	7	ADP31605	Adp31605 Human SAR
717	16	61.5	71	4	AAH27022	Aah27022 Peptide #	790	16	61.5	152	8	ADV95514	Adv95514 Human pol
718	16	61.5	71	4	ABM28383	Abm28383 Peptide #	791	16	61.5	153	8	ABP62822	Abp62822 Human CGD
719	16	61.5	71	4	ABM19018	Abm19018 Protein #	792	16	61.5	154	5	AAH36181	Aah36181 Human CGD
720	16	61.5	71	4	AAH67735	Aah67735 Human bon	793	16	61.5	154	6	AAH36181	Aah36181 Human CGD
721	16	61.5	71	4	AAU53020	Aau53020 Propionib	794	16	61.5	155	7	ABO77012	AbO77012 Pseudomon
722	16	61.5	71	4	AAH54335	Aah54335 Human bra	795	16	61.5	156	7	ABO83103	AbO83103 Pseudomon
723	16	61.5	71	4	ABG48403	Abg48403 Human liv	796	16	61.5	156	7	ABO76338	AbO76338 Pseudomon
724	16	61.5	71	5	ABM02326	Abm02326 Peptide #	797	16	61.5	158	7	ABO79744	AbO79744 Human gen
725	16	61.5	71	5	ABG53389	Abg53389 Human pep	798	16	61.5	158	8	ABO59415	AbO59415 Human gen
726	16	61.5	71	6	ABM49539	Abm49539 Propionib	799	16	61.5	159	4	AAO10454	Aao10454 Human pol
727	16	61.5	71	9	ADZ47683	Adz47683 Ant1-1L-1	800	16	61.5	159	7	ABO79796	AbO79796 Pseudomon
728	16	61.5	74	4	AAU47686	Aau47686 Propionib	801	16	61.5	160	7	ABO73965	AbO73965 Pseudomon
729	16	61.5	74	6	ABM44205	Abm44205 Human bon	802	16	61.5	162	3	AAH95932	Aah95932 Porcine a
730	16	61.5	80	4	AAH70800	Aah70800 Human bon	803	16	61.5	162	7	ABO69461	AbO69461 Pseudomon
731	16	61.5	80	4	AAH59526	Aah59526 Propionib	804	16	61.5	162	7	ABO79469	AbO79469 Pseudomon
732	16	61.5	80	6	ABM56045	Abm56045 Human col	805	16	61.5	162	8	ADP30813	Adp30813 Human sec
733	16	61.5	88	4	AAH74029	Aah74029 Human sec	806	16	61.5	162	8	ADH88496	Adh88496 Plant ful
734	16	61.5	90	8	ADP31399	Adp31399 Human sec	807	16	61.5	165	4	AAU54621	Aau54621 Propionib
735	16	61.5	90	8	ADP31400	Adp31400 Human sec	808	16	61.5	165	6	ABM51140	Abm51140 Propionib
736	16	61.5	92	4	AAU46180	Aau46180 Propionib	809	16	61.5	165	8	ADP31286	Adp31286 Human sec
737	16	61.5	92	6	ABM43699	Abm43699 Propionib	810	16	61.5	167	2	AAH38569	Aah38569 Neisseria
738	16	61.5	93	4	ABM43737	Abm43737 Peptide #	811	16	61.5	167	2	AAH38569	Aah38569 Neisseria
739	16	61.5	93	4	AAH37635	Aah37635 Peptide #	812	16	61.5	167	9	ABH84898	Abh84898 N. mening
740	16	61.5	93	4	AAH37635	Aah37635 Peptide #	813	16	61.5	167	9	ABH84898	Abh84898 N. mening
741	16	61.5	93	5	ABG46494	Abg46494 Human pep	814	16	61.5	168	7	ABO70841	AbO70841 Human sec
742	16	61.5	93	5	ADP30859	Adp30859 Human sec	815	16	61.5	168	8	ADP30841	Adp30841 Human gen
743	16	61.5	95	8	ADG23362	Adg22362 Cyanophag	816	16	61.5	171	8	ADP30599	Adp30599 Human sec
744	16	61.5	104	2	AAH12843	Aah12843 HTLV-1 pr	817	16	61.5	171	8	ABO74910	AbO74910 Pseudomon
745	16	61.5	104	2	AAH12843	Aah12843 HTLV-1 pr	818	16	61.5	174	7	ABO74910	AbO74910 Pseudomon
746	16	61.5	104	6	AAU55097	Aau55097 Propionib	819	16	61.5	174	8	ADP30801	Adp30801 Human sec
747	16	61.5	104	6	ABM51616	Abm51616 Propionib	820	16	61.5	175	4	ABH86803	Abh86803 Drosophil
748	16	61.5	104	7	ABO78018	AbO78018 Pseudomon	821	16	61.5	180	8	ADP30820	Adp30820 Human sec
749	16	61.5	105	4	AAU42091	Aau42091 Propionib	822	16	61.5	180	8	ADP30825	Adp30825 Human sec
750	16	61.5	105	6	ABM38610	Abm38610 Propionib	823	16	61.5	180	8	ADP30821	Adp30821 Human sec
751	16	61.5	106	6	ABM55201	Abm55201 Propionib	824	16	61.5	180	8	ADP30825	Adp30825 Human sec
752	16	61.5	108	4	AAU42243	Aau42243 Propionib	825	16	61.5	180	8	ADP30827	Adp30827 Human sec
753	16	61.5	108	6	ABM38762	Abm38762 Propionib	826	16	61.5	181	8	ADP31029	Adp31029 Human sec
754	16	61.5	108	6	ABM38762	Abm38762 Propionib	827	16	61.5	181	8	ADP31029	Adp31029 Human sec

828	16	61.5	181	8	ADQ67035	Adq67035	Novel hum	901	16	61.5	266	9	AEb42329	Aeb42329	L. pneumo
829	16	61.5	185	4	ABb11163	Abb11163	Human ctp	902	16	61.5	267	8	ADp30822	Adp30822	Human sec
830	16	61.5	186	7	ABo68891	AbO68891	Pseudomon	903	16	61.5	267	8	ADp31483	Adp31483	Human sec
831	16	61.5	186	8	ADp30789	Adp30789	Human sec	904	16	61.5	268	8	ADp31697	Adp31697	Human sec
832	16	61.5	186	8	ADp30787	Adp30787	Human sec	905	16	61.5	270	8	ADp31435	Adp31435	Human sec
833	16	61.5	186	8	ADp30786	Adp30786	Human sec	906	16	61.5	272	5	AAO20952	Aao20952	Human-N-K
834	16	61.5	186	8	ADp30785	Adp30785	Human sec	907	16	61.5	272	6	ABg76399	Abg76399	Antno act
835	16	61.5	187	7	ABO68818	AbO68818	Pseudomon	908	16	61.5	273	8	ADp31336	Adp31336	Human sec
836	16	61.5	188	8	ADp96439	Adp96439	Plant ful	909	16	61.5	273	8	ADp31583	Adp31583	Human sec
837	16	61.5	189	8	ADp30641	Adp30641	Human sec	910	16	61.5	274	8	ADp31679	Adp31679	Human sec
838	16	61.5	192	8	ADp31335	Adp31335	Human sec	911	16	61.5	275	8	ADp65898	Adp65898	Novel hum
839	16	61.5	192	8	ADp31576	Adp31576	Human sec	912	16	61.5	275	8	ADp60000	Adp60000	Plant pol
840	16	61.5	193	7	ABO73977	AbO73977	Pseudomon	913	16	61.5	277	8	ADp30503	Adp30503	Human sec
841	16	61.5	195	7	ADp30696	Adp30696	Human sec	914	16	61.5	278	7	AAO66278	Aao66278	Arabiidops
842	16	61.5	197	7	ABO73630	AbO73630	Pseudomon	915	16	61.5	278	7	ABO68255	AbO68255	Pseudomon
843	16	61.5	198	8	ADp30477	Adp30477	Human sec	916	16	61.5	279	3	AAO32271	Aao32271	Arabiidops
844	16	61.5	198	8	ADp30481	Adp30481	Human sec	917	16	61.5	279	3	AAO64649	Aao64649	Arabiidops
845	16	61.5	198	8	ADp30840	Adp30840	Human sec	918	16	61.5	279	8	ADp31553	Adp31553	Human sec
846	16	61.5	200	3	AAO54845	Aao54845	Arabiidops	919	16	61.5	279	8	ADp31107	Adp31107	Human sec
847	16	61.5	201	8	ADp30711	Adp30711	Human sec	920	16	61.5	280	8	ADp66570	Adp66570	Novel hum
848	16	61.5	201	8	ADp31351	Adp31351	Human sec	921	16	61.5	280	8	ADp09856	Adp09856	Human pro
849	16	61.5	203	8	ADp66375	Adp66375	Novel hum	922	16	61.5	282	7	ADp92143	Adp92143	Human hai
850	16	61.5	204	8	ADp31404	Adp31404	Human sec	923	16	61.5	285	3	AAO68088	Aao68088	Arabiidops
851	16	61.5	209	4	ABb70477	Abb70477	Drosoph11	924	16	61.5	286	7	ABO66150	AbO66150	Klebsiell
852	16	61.5	210	8	ADp58395	Adp58395	Plant pol	925	16	61.5	288	8	ADp31463	Adp31463	Human sec
853	16	61.5	210	8	ADp66008	Adp66008	Plant ful	926	16	61.5	291	8	ADp31193	Adp31193	Human sec
854	16	61.5	212	7	ABO79141	AbO79141	Pseudomon	927	16	61.5	297	8	ADp30744	Adp30744	Human sec
855	16	61.5	213	8	ADp31339	Adp31339	Human sec	928	16	61.5	298	8	ADp43100	Adp43100	Plant tra
856	16	61.5	214	8	ABM64290	AbM64290	Human sec	929	16	61.5	300	8	ADp30616	Adp30616	Human sec
857	16	61.5	215	7	ABE08255	AbE08255	Novel pro	930	16	61.5	300	8	ADp30775	Adp30775	Human sec
858	16	61.5	216	4	AAO75257	Aao75257	Human col	931	16	61.5	309	2	AAp37650	Aap37650	Chlamydia
859	16	61.5	217	8	ADp10185	Adp10185	Plant ful	932	16	61.5	310	8	ADp30521	Adp30521	Human sec
860	16	61.5	218	7	ABO67270	AbO67270	Klebsiell	933	16	61.5	311	3	AAO66277	Aao66277	Arabiidops
861	16	61.5	219	8	ADp30593	Adp30593	Human sec	934	16	61.5	312	3	AAO64648	Aao64648	Arabiidops
862	16	61.5	221	4	AAO30388	Aao30388	C glutami	935	16	61.5	312	3	AAO32270	Aao32270	Arabiidops
863	16	61.5	222	8	ADp31379	Adp31379	Human sec	936	16	61.5	316	3	AAO66276	Aao66276	Arabiidops
864	16	61.5	222	8	ADp31300	Adp31300	Human sec	937	16	61.5	317	3	AAO64647	Aao64647	Arabiidops
865	16	61.5	223	7	ABO76890	AbO76890	Pseudomon	938	16	61.5	317	3	AAO32269	Aao32269	Arabiidops
866	16	61.5	226	7	ABO75147	AbO75147	Pseudomon	939	16	61.5	317	5	AAU92968	Aau92968	Arabiidops
867	16	61.5	228	8	ADp31577	Adp31577	Human sec	940	16	61.5	317	7	ADp30646	Adp30646	Plant yie
868	16	61.5	228	8	ADp31334	Adp31334	Human sec	941	16	61.5	317	8	ADp41555	Adp41555	Plant tra
869	16	61.5	228	8	ADp30764	Adp30764	Human sec	942	16	61.5	317	8	ADp31615	Adp31615	Human sec
870	16	61.5	229	4	AAO91496	Aao91496	C glutami	943	16	61.5	317	8	ADp06984	Adp06984	Arabiidops
871	16	61.5	229	6	ABU70511	Abu70511	Human adi	944	16	61.5	317	9	AEA26635	Aea26635	Streess to
872	16	61.5	231	5	ABJ01955	Abj01955	158PID7 S	945	16	61.5	318	3	AAO68087	Aao68087	Arabiidops
873	16	61.5	231	8	ADp30610	Adp30610	Human sec	946	16	61.5	318	8	ADp30588	Adp30588	Human sec
874	16	61.5	233	3	AAp74791	Aap74791	Neisseria	947	16	61.5	320	8	ADp31607	Adp31607	Human sec
875	16	61.5	233	4	ADp788408	Adp788408	Novel SAR	948	16	61.5	320	8	ADp31649	Adp31649	Human sec
876	16	61.5	234	4	ABG10334	Abg10334	Novel hum	949	16	61.5	320	8	ABM62688	Abm62688	Human dia
877	16	61.5	234	8	ADp30484	Adp30484	Human sec	950	16	61.5	323	3	AAO68086	Aao68086	Arabiidops
878	16	61.5	234	8	ADp31166	Adp31166	Human sec	951	16	61.5	323	4	AAAB84370	Aaab84370	Antno act
879	16	61.5	234	8	ADp31533	Adp31533	Plant ful	952	16	61.5	326	7	ABO66821	AbO66821	Klebsiell
880	16	61.5	236	8	ABM64301	Abm64301	Human dia	953	16	61.5	327	8	ADp30478	Adp30478	Human sec
881	16	61.5	236	8	ABM64275	Abm64275	Human dia	954	16	61.5	327	8	ADp31660	Adp31660	Plant ful
882	16	61.5	236	8	ABM64286	Abm64286	Human dia	955	16	61.5	328	7	ABO77840	AbO77840	Pseudomon
883	16	61.5	240	7	ABO73245	AbO73245	Pseudomon	956	16	61.5	339	8	ADp30892	Adp30892	Human sec
884	16	61.5	240	8	ADp31579	Adp31579	Human sec	957	16	61.5	341	7	ABO66273	AbO66273	Human sec
885	16	61.5	240	8	ADp31475	Adp31475	Human sec	958	16	61.5	344	3	AAp75045	Aap75045	Neisseria
886	16	61.5	243	8	ADp30983	Adp30983	Human sec	959	16	61.5	345	8	ADp31204	Adp31204	Human sec
887	16	61.5	243	8	ADp31343	Adp31343	Human sec	960	16	61.5	346	8	ADp31355	Adp31355	Human sec
888	16	61.5	246	7	ABO74848	AbO74848	Pseudomon	961	16	61.5	348	4	ABBE8968	Abbe8968	Drosoph11
889	16	61.5	246	7	ABO75282	AbO75282	Pseudomon	962	16	61.5	348	8	ADp34950	Adp34950	Xylanase
890	16	61.5	249	7	ABO680226	AbO680226	Pseudomon	963	16	61.5	348	8	ADp30985	Adp30985	Human sec
891	16	61.5	249	8	ADp30645	Adp30645	Human sec	964	16	61.5	348	8	ADp31676	Adp31676	Human sec
892	16	61.5	249	8	ADp30754	Adp30754	Human sec	965	16	61.5	349	8	ADp48010	Adp48010	Polypepti
893	16	61.5	252	4	ABBE3479	Abbe3479	Drosoph11	966	16	61.5	350	8	ADp31672	Adp31672	Human sec
894	16	61.5	253	4	ABO70290	AbO70290	Pseudomon	967	16	61.5	353	8	ADp31558	Adp31558	Human sec
895	16	61.5	254	8	ADp31396	Adp31396	Human sec	968	16	61.5	356	4	ABG04360	Abg04360	Novel hum
896	16	61.5	255	8	ADp30658	Adp30658	Human sec	969	16	61.5	358	8	ADp31063	Adp31063	Human sec
897	16	61.5	252	8	ADp31674	Adp31674	Human sec	970	16	61.5	363	7	ABO68771	AbO68771	Pseudomon
898	16	61.5	263	8	ADp31401	Adp31401	Human sec	971	16	61.5	363	8	ADp30659	Adp30659	Human sec
899	16	61.5	264	8	ADp31456	Adp31456	Human sec	972	16	61.5	366	8	ADp31106	Adp31106	Human sec
900	16	61.5	266	9	AEb37222	Aeb37222	L. pneumo	973	16	61.5	366	8	ADp31199	Adp31199	Human sec

974	16	61.5	366	8	ADP31075	Adp31075 Human sec
975	16	61.5	369	7	ABO72044	AbO72044 Pseudomon
976	16	61.5	369	8	ADP31289	Adp31289 Human sec
977	16	61.5	372	8	ADP31610	Adp31610 Human sec
978	16	61.5	375	9	ADX06695	Adx06695 Cyc1n-de
979	16	61.5	380	3	AAG04919	Aag04919 Arabidops
980	16	61.5	380	3	AAG32553	Aag32553 Arabidops
981	16	61.5	381	2	AAV38570	Aav38570 Neisseria
982	16	61.5	381	2	AAV75046	Aav75046 Neisseria
983	16	61.5	381	9	ABE48989	Aeb48989 N. mening
984	16	61.5	382	2	AAV38571	Aav38571 Neisseria
985	16	61.5	382	3	AAV75044	Aav75044 Neisseria
986	16	61.5	382	4	ABB70181	Abb70181 Drosophil
987	16	61.5	382	9	ABE48990	Aeb48990 N. gonorr
988	16	61.5	384	8	ADP31338	Adp31338 Human sec
989	16	61.5	384	8	ADP31318	Adp31318 Human sec
990	16	61.5	387	8	ADP31158	Adp31158 Human sec
991	16	61.5	387	8	ADZ47819	Adz47819 Amino aci
992	16	61.5	389	8	ADX78458	Adx78458 Plant ful
993	16	61.5	389	8	ADY24278	Ady24278 Plant ful
994	16	61.5	391	7	ADN31030	Adn31030 Fruit fly
995	16	61.5	396	8	ADP31625	Adp31625 Human sec
996	16	61.5	396	8	ADP31378	Adp31378 Human sec
997	16	61.5	398	4	AAM40928	Aam40928 Human pol
998	16	61.5	398	7	ADF60215	Adf60215 Human con
999	16	61.5	398	7	ABO71161	AbO71161 Pseudomon
1000	16	61.5	398	8	ADP31604	Adp31604 Human sec

ALIGNMENTS

```
RESULT 1
ID ADZ47471 standard; peptide; 30 AA.
AC
XX ADZ47471;
XX
DT 30-JUN-2005 (first entry)
XX
DE Alu sense peptide SEQ ID NO 50.
XX
XX hyaluronan synthase; HAS; hyaluronidase; HYAL; HAS1; HAS2; HAS3; HYAL1;
KM HYAL2; HAL3; PH-20; cancer; hyperproliferation; inflammation; anabolic;
KM hypertensive; gynaecological; neuroprotective; antianemic; cycostatic;
KM anti-inflammatory; endocrine-gen; immunosuppressive; gene therapy.
XX
OS Unidentified.
XX
XX WO2005035548-A1.
XX
XX 21-APR-2005.
XX
XX 11-OCT-2004; 2004WO-AU001383.
XX
XX 10-OCT-2003; 2003AU-00905551.
XX
XX 01-DEC-2003; 2003AU-00906658.
XX
XX (MEDI-) MEDITECH RES LTD.
XX
XX Brown TJ, Brownlee GR;
XX
XX WPI; 2005-315540/32.
XX
XX Novel compound capable of reducing level, function or activity of
PT hyaluronan synthase or hyaluronidase, useful for treating cancer,
PT hyperproliferative condition, A-beta-lipoproteinemia, A-V, A beta-2-
PT microglobulin amyloidosis.
XX
XX Example 22; SEQ ID NO 50; 277pp; English.
XX
XX This invention describes a novel nucleic acid molecule capable of
CC reducing the level of hyaluronan synthase (HAS) or hyaluronidase (HYAL)
```

or the function or activity of HAS or HYAL. HAS is selected from HAS1, HAS2 and HAS3. HYAL is chosen from HYAL1, HYAL2, HAL3 and PH-20. The nucleic acid molecule is an oligonucleotide or its chemically modified form comprising a chemically modified backbone or a non-natural internucleoside linkage. The compound is an interactive molecule capable of binding or otherwise associating with HAS and/or HYAL to reduce HAS and/or HYAL function or activity. The novel molecule is useful for treating cancer, hyperproliferative conditions or inflammatory conditions. The products of the invention are useful for treating or prophylaxis of a condition in subject, which involves administering to the subject, an HAS and/or HYAL level-reducing or HAS and/or HYAL activity reducing effective amount of active molecule. The products of the invention have anabolic, hypertensive, gynaecological, neuroprotective, antianemic, cycostatic, anti-inflammatory, endocrine-gen, and immunosuppressive activity. This sequence represents a peptide used in the method of the invention. NOTE: The specification describes SEQ ID NO 27-51 as nucleotide primers however the Sequence Listing represents the nucleotides in a three letter amino acid code. The nucleotide sequences have been made and are represented in { }.

SQ Sequence 30 AA;

Query Match 69.2%; Score 18; DB 9; Length 30;

Best Local Similarity 20.0%; Pred. No. 2.4;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

DY 1 CXXXXXXXXX 10
DB 18 CTAATAATAC 27

RESULT 2

```
ADP30533
ID ADP30533 standard; protein; 89 AA.
AC
XX ADP30533;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1300.
XX
XX Cycostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX
XX 29-AUG-2002; 2002US-0406579P.
XX
XX 29-AUG-2002; 2002US-0406585P.
XX
XX 29-AUG-2002; 2002US-0406588P.
XX
XX 29-AUG-2002; 2002US-0406608P.
XX
XX 29-AUG-2002; 2002US-0406611P.
XX
XX 29-AUG-2002; 2002US-0406612P.
XX
XX 29-AUG-2002; 2002US-0406616P.
XX
XX 29-AUG-2002; 2002US-0406640P.
XX
XX 29-AUG-2002; 2002US-0406642P.
XX
XX 29-AUG-2002; 2002US-0406646P.
XX
XX 29-AUG-2002; 2002US-0406653P.
XX
XX 29-AUG-2002; 2002US-0406655P.
XX
XX 29-AUG-2002; 2002US-0406666P.
XX
XX 17-SEP-2002; 2002US-0410946P.
XX
XX 17-SEP-2002; 2002US-0410947P.
XX
XX 17-SEP-2002; 2002US-0410948P.
XX
XX 17-SEP-2002; 2002US-0410949P.
XX
XX 17-SEP-2002; 2002US-0410953P.
XX
XX 17-SEP-2002; 2002US-0410957P.
XX
XX 17-SEP-2002; 2002US-0410958P.
```

(FIVE-) FIVE PRIME THERAPEUTICS INC.

William LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
Hallenbeck RF, Huang MM, Kochakota S, Haishan L, Linemann T;
Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

WPI, 2004-348438/32.

New nucleic acid molecule for diagnosing, preventing or treating diseases such as proliferative (e.g. cancer), inflammatory, immune, metabolic, genetic, bacterial and viral diseases.

Claim 1; SEQ ID NO 2531; 428BP; English.

The present invention relates to an isolated nucleic acid molecule encoding a polypeptide which is believed to be cytostatic, anti-inflammatory, immunosuppressive, antibacterial and antiviral. The composition and methods are useful for diagnosing, preventing and treating diseases such as proliferative (e.g. cancer), inflammatory, immune, metabolic, genetic, bacterial and viral diseases. The present sequence represents a human secreted protein. The present sequence is available on WIPOMB and is not in the specification.

Sequence 69 AA:

Query Match	Score	DB	Length
Best Local Similarity	20.0%	Pred. No. 3.1	
Matches	2	Conservative	0
		Mismatches	8
		Indels	0
		Gaps	0

QY	1	XXXXXXXXXX	10
Db	24	CTTAAAAAAC	33

RESULT 3
ADP30531
ID ADP30531 standard; protein; 89 AA.
XX AC ADP30531;
XX DT 12-AUG-2004 (first entry)
XX DE Human secreted protein SEQ ID #1298.
XX KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX KW cancer; inflammatory; immune; human secreted protein.
XX OS Homo sapiens.
XX WO2004035732-A2.
XX PD 29-APR-2004.
XX PF 28-AUG-2003; 2003WO-US026780.
XX PR 29-AUG-2002; 2002US-0406576P.
XX PR 29-AUG-2002; 2002US-0406579P.
XX PR 29-AUG-2002; 2002US-0406585P.
XX PR 29-AUG-2002; 2002US-0406588P.
XX PR 29-AUG-2002; 2002US-0406608P.
XX PR 29-AUG-2002; 2002US-0406611P.
XX PR 29-AUG-2002; 2002US-0406612P.
XX PR 29-AUG-2002; 2002US-0406616P.
XX PR 29-AUG-2002; 2002US-0406640P.
XX PR 29-AUG-2002; 2002US-0406642P.
XX PR 29-AUG-2002; 2002US-0406646P.
XX PR 29-AUG-2002; 2002US-0406653P.
XX PR 29-AUG-2002; 2002US-0406655P.
XX PR 29-AUG-2002; 2002US-0406666P.
XX PR 17-SEP-2002; 2002US-0410946P.
XX PR 17-SEP-2002; 2002US-0410947P.
XX PR 17-SEP-2002; 2002US-0410948P.
XX PR 17-SEP-2002; 2002US-0410949P.
XX PR 17-SEP-2002; 2002US-0410953P.
XX PR 17-SEP-2002; 2002US-0410957P.
XX PR 17-SEP-2002; 2002US-0410958P.
XX PR 17-SEP-2002; 2002US-0410959P.
XX PR 17-SEP-2002; 2002US-0410961P.
XX PR 17-SEP-2002; 2002US-0410962P.
XX PR 17-SEP-2002; 2002US-0411019P.
XX PR 17-SEP-2002; 2002US-0411022P.
XX PR 17-SEP-2002; 2002US-0411023P.
XX PR 17-SEP-2002; 2002US-0411024P.
XX PR 17-SEP-2002; 2002US-0411032P.
XX PR 17-SEP-2002; 2002US-0411035P.
XX PR 17-SEP-2002; 2002US-0411037P.
XX PR 17-SEP-2002; 2002US-0411041P.
XX PR 17-SEP-2002; 2002US-0411045P.
XX PR 17-SEP-2002; 2002US-0411046P.
XX PR 17-SEP-2002; 2002US-0411048P.
XX PR 17-SEP-2002; 2002US-0411052P.
XX PR 17-SEP-2002; 2002US-0411055P.
XX PR 17-SEP-2002; 2002US-0411073P.
XX PR 17-SEP-2002; 2002US-0411082P.
XX PR 17-SEP-2002; 2002US-0411101P.
XX PR 17-SEP-2002; 2002US-0411111P.
XX PR 18-APR-2003; 2003US-0463700P.

```

PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485244P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486981P.
PR 15-JUL-2003; 2003US-0486980P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
XX Halebeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
XX Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2529; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cytosstatic,
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPOWEB and is not in the specification.
XX
XX Sequence 89 AA;
SQ
Query Match 69.2%; Score 18; DB 8; Length 89;
Best Local Similarity 20.0%; Pred. No. 3.1;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXXX 10
Db 24 CTTAAAAAAC 33
RESULT 4
AAM93527
ID AAM93527 standard; protein; 109 AA.
XX
XX AAM93527;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polypeptide, SEQ ID NO: 3263.
XX
XX Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
XX Homo sapiens.

```

```

XX
XX EP130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-00114089.
XX
XX 08-JUL-1999; 99JP-00194486.
XX 11-JAN-2000; 2000JP-00118774.
XX 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Oca T, Nishikawa T, Isogai T, Hayaishi K, Ishii S, Kawai Y;
XX Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
XX
XX N-PSDB; AAK94457.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their use
XX PT in genetic manipulation.
XX
XX Claim 8; SEQ ID NO 3263; 1380bp + Sequence listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
XX CC clones. 830 cDNA molecules encoding a human protein have been isolated
XX CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX CC been determined. Primers for synthesizing the full length cDNA are useful
XX CC for clarifying the function of the protein encoded by the cDNA. The full
XX CC length clones were obtained by construction of full length enriched cDNA
XX CC libraries that were synthesized by the oligo-capping method. The primers
XX CC enable the production of the full length cDNA easily without any special
XX CC methods. The present sequence is a polypeptide encoded by a full length
XX CC human cDNA of the invention. Note: The sequence data for this patent did
XX CC not form part of the printed specification, but was obtained in CD-ROM
XX CC format directly from EPO
XX
XX Sequence 109 AA;
SQ
Query Match 69.2%; Score 18; DB 4; Length 109;
Best Local Similarity 20.0%; Pred. No. 3.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXXX 10
Db 4 CSSSARSTGC 13
RESULT 5
ADL31230
ID ADL31230 standard; protein; 109 AA.
XX
XX ADL31230;
XX
XX 20-MAY-2004 (first entry)
XX
XX Human protein encoded by a full length cDNA clone Segid 3263.
XX
XX human; medicine; signal transduction; glycoprotein; transcription;
XX KW oligo-capping method.
XX
XX Homo sapiens.
XX
XX EP1396543-A2.
XX
XX 10-MAR-2004.
XX
XX 07-JUL-2000; 2003EP-00025638.
XX
XX 08-JUL-1999; 99JP-00194486.
XX 11-JAN-2000; 2000JP-00118774.
XX 02-MAY-2000; 2000JP-00183865.
XX 07-JUL-2000; 2000EP-00114089.

```

XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 PA Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2004-204755/20.
 DR N-PSDB; ADL31229.
 XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full
 PT length human cDNAs.
 PS Example 1; SEQ ID NO 3263; 1340bp; English.
 XX This invention relates to a novel primers useful for synthesizing full
 CC length cDNA molecules that encode human proteins. Specifically, it refers
 CC to secretory or membrane proteins that are potential therapeutic agents/
 CC target molecules in the field of medicine, and in particular genes
 CC encoding proteins that are associated with signal transduction,
 CC glycoproteins and transcription. The present invention describes a method
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'
 CC ends using the oligo-capping method. This polypeptide sequence is a full
 CC length human protein of the invention.
 CC
 SQ Sequence 109 AA;
 Query Match 69.2%; Score 18; DB 8; Length 109;
 Best Local Similarity 20.0%; Pred. No. 3.2;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 Oy 1 CXXXXXXXXXC 10
 Db 4 CSSSATSTSC 13
 RESULT 6
 ADP30690
 ID ADP30690 standard; protein; 109 AA.
 AC ADP30690;
 DT 12-AUG-2004 (first entry)
 XX
 DE Human secreted protein SEQ ID #1457.
 XX
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
 KM cancer; inflammatory; immune; human secreted protein.
 XX
 OS Homo sapiens.
 PN WO2004035732-A2.
 XX
 PD 29-APR-2004.
 XX
 PF 28-AUG-2003; 2003WO-US026780.
 XX
 PR 29-AUG-2002; 2002US-0406576P.
 PR 29-AUG-2002; 2002US-0406579P.
 PR 29-AUG-2002; 2002US-0406585P.
 PR 29-AUG-2002; 2002US-0406588P.
 PR 29-AUG-2002; 2002US-0406608P.
 PR 29-AUG-2002; 2002US-0406611P.
 PR 29-AUG-2002; 2002US-0406612P.
 PR 29-AUG-2002; 2002US-0406616P.
 PR 29-AUG-2002; 2002US-0406640P.
 PR 29-AUG-2002; 2002US-0406642P.
 PR 29-AUG-2002; 2002US-0406646P.
 PR 29-AUG-2002; 2002US-0406653P.
 PR 29-AUG-2002; 2002US-0406655P.
 PR 29-AUG-2002; 2002US-0406666P.
 PR 17-SEP-2002; 2002US-0410946P.
 PR 17-SEP-2002; 2002US-0410947P.
 PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.
 PR 17-SEP-2002; 2002US-0410953P.
 PR 17-SEP-2002; 2002US-0410957P.
 PR 17-SEP-2002; 2002US-0410958P.
 PR 17-SEP-2002; 2002US-0410959P.
 PR 17-SEP-2002; 2002US-0410960P.
 PR 17-SEP-2002; 2002US-0410961P.
 PR 17-SEP-2002; 2002US-0410962P.
 PR 17-SEP-2002; 2002US-0411019P.
 PR 17-SEP-2002; 2002US-0411022P.
 PR 17-SEP-2002; 2002US-0411023P.
 PR 17-SEP-2002; 2002US-0411024P.
 PR 17-SEP-2002; 2002US-0411032P.
 PR 17-SEP-2002; 2002US-0411035P.
 PR 17-SEP-2002; 2002US-0411037P.
 PR 17-SEP-2002; 2002US-0411041P.
 PR 17-SEP-2002; 2002US-0411045P.
 PR 17-SEP-2002; 2002US-0411046P.
 PR 17-SEP-2002; 2002US-0411048P.
 PR 17-SEP-2002; 2002US-0411052P.
 PR 17-SEP-2002; 2002US-0411055P.
 PR 17-SEP-2002; 2002US-0411073P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411101P.
 PR 17-SEP-2002; 2002US-0411111P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463709P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 02-MAY-2003; 2003US-0467230P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 19-MAY-2003; 2003US-0471336P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 22-MAY-2003; 2003US-0472430P.
 PR 09-JUN-2003; 2003US-0476609P.
 PR 09-JUN-2003; 2003US-0476641P.
 PR 08-JUL-2003; 2003US-0485218P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485224P.
 PR 08-JUL-2003; 2003US-0485255P.
 PR 14-JUL-2003; 2003US-0486446P.
 PR 14-JUL-2003; 2003US-048680P.
 PR 15-JUL-2003; 2003US-0486891P.
 PR 15-JUL-2003; 2003US-0486960P.
 PR 08-AUG-2003; 2003US-0493341P.
 PR 08-AUG-2003; 2003US-0493370P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.
 XX
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
 XX
 PI Williams LT, Chu K, Lee E, Healtir K, Beaurang PA, Behrens D;
 PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
 DR WPI; 2004-348438/32.
 XX
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 PT genetic, bacterial and viral diseases.
 PS Claim 1; SEQ ID NO 2688; 428bp; English.
 XX
 CC The present invention relates to an isolated nucleic acid molecule
 CC encoding a polypeptide which is believed to be cytostatic,
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is

CC available on WIPOMEB and is not in the specification.
XX Sequence 109 AA;
SQ

Query Match 69.2%; Score 18; DB 8; Length 109;
Best Local Similarity 20.0%; Pred. No. 3.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXKC 10
DB 96 CTTTATATC 105

RESULT 7
ID ADP31474 standard; protein; 144 AA.
XX ADP31474;
AC
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2241.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
XX MO2004035732-A2.
XX
XX 29-APR-2004.
PD
XX 28-AUG-2003; 2003MO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-041101P.
PR 17-SEP-2002; 2002US-041111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisnan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
DR
XX
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
XX Claim 1; SEQ ID NO 3472; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
XX
SQ Sequence 144 AA;
QY 1 CXXXXXXKC 10
DB 72 CAATTAATC 81

Query Match 69.2%; Score 18; DB 8; Length 144;
Best Local Similarity 20.0%; Pred. No. 3.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

RESULT 8
AB069151
ID AB069151 standard; protein; 167 AA.
XX
XX AB069151;
AC
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #1326.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX Pseudomonas aeruginosa.
 OS
 XX
 PN US6551795-B1.
 XX
 PD 22-APR-2003.
 XX
 PF 18-FEB-1999; 99US-00252991.
 XX
 PR 18-FEB-1998; 98US-0074788P.
 PR 27-JUL-1998; 98US-0094190P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 DR WPI; 2003-615309/58.
 DR N-PSDB; ABD02722.
 XX
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 PS
 PS Disclosure; SEQ ID NO 17897; 455bp; English.
 CC The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biotech technology. Sequences ABO67826-
 CC ABO68396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 CC
 SQ Sequence 167 AA;
 Query Match 69.2%; Score 18; DB 7; Length 167;
 Best Local Similarity 20.0%; Pred. No. 3.5;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 CXXXXXXXXX 10
 DB 60 CSTSTSTTSC 69
 RESULT 9
 ADS12185
 ID ADS12185 standard; protein; 169 AA.
 XX
 AC ADS12185;
 XX
 DT 16-DEC-2004 (first entry)
 XX
 DE Human therapeutic contig protein - SEQ ID 2422.
 XX
 KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;
 KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
 KW aplastic anaemia; cancer; wound healing; gene therapy.
 OS
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1.169

FT /label= Unknown, OTHER
 FT /note= "OTHER = In-frame STOP codon"
 XX
 PN WO2004080148-A2.
 XX
 PD 23-SEP-2004.
 XX
 PF 30-SEP-2003; 2003WO-US030720.
 XX
 PR 02-OCT-2002; 2002US-0416186P.
 XX
 PA (NUVE-) NUVELO INC.
 XX
 PI Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y,
 PI Wang D, Chen R, Zhao Q, Wang J, Ghosh M, Xue AJ, Wang G, Zhou P;
 DR WPI; 2004-668857/65.
 DR N-PSDB; ADS11587.
 XX
 PT New polynucleotide, useful in preparing a composition for diagnosing or
 PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
 PT aplastic anemia or cancer for promoting wound healing.
 PS
 PS Example 2; SEQ ID NO 2422; 718bp; English.
 CC
 CC The invention relates to a novel isolated polynucleotide and the encoded
 CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
 CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may
 CC be useful in preparing a composition for diagnosing or treating
 CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
 CC disorders, such as aplastic anaemia or cancer, as well as for promoting
 CC wound healing. The molecules may also be utilised during gene therapy
 CC procedures. The current sequence is that of a human therapeutic contig
 CC protein of the invention.
 CC
 SQ Sequence 169 AA;
 Query Match 69.2%; Score 18; DB 8; Length 169;
 Best Local Similarity 20.0%; Pred. No. 3.5;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 CXXXXXXXXX 10
 DB 81 CAAAAAAAC 90
 RESULT 10
 ADP30794
 ID ADP30794 standard; protein; 171 AA.
 XX
 AC ADP30794;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human secreted protein SEQ ID #1561.
 XX
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
 KW cancer; inflammatory; immune; human secreted protein.
 OS
 XX Homo sapiens.
 XX
 PN WO2004035732-A2.
 XX
 PD 29-APR-2004.
 XX
 PF 28-AUG-2003; 2003WO-US026780.
 XX
 PR 29-AUG-2002; 2002US-0406576P.
 PR 29-AUG-2002; 2002US-0406579P.
 PR 29-AUG-2002; 2002US-0406585P.
 PR 29-AUG-2002; 2002US-0406588P.
 PR 29-AUG-2002; 2002US-0406608P.
 PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 29-AUG-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485242P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Heestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MW, Kothakota S, Hsieh L, Linemann T;
XX Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
such as proliferative (e.g. cancer), inflammatory, immune, metabolic,

PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2792; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antitumorigenic, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 171 AA;
Query Match 69.2%; Score 18; DB 8; Length 171;
Best Local Similarity 20.0%; Pred. No. 3.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXC 10
Db 146 CAATTAAC 155
RESULT 11
ADP30793
ID ADP30793 standard; protein; 171 AA.
XX
AC ADP30793;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1560.
XX
KW Cytostatic; Antitumorigenic; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.

PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411010P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485255P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX
DR WPI; 2004-348438/32.
XX
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 2791; 428bp; English.
XX
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic.
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
XX
SQ Sequence 171 AA;

Query Match 69.2%; Score 18; DB 8; Length 171;
Best Local Similarity 20.0%; Pred. No. 3.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
1 CXXXXXXXXX 10
146 CAATATAAAC 155

RESULT 12
ADP30806
ID ADP30806 standard; protein; 183 AA.
XX
XX
AC ADP30806;
XX
XX
DT 12-AUG-2004 (first entry)
XX
XX
DE Human secreted protein SEQ ID #1573.
XX
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
XX
PN WO2004035732-A2.
XX
XX
PD 29-APR-2004.
XX
XX
PF 28-AUG-2003; 2003MO-US026780.
XX
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411012P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 14-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX Claim 1; SEQ ID NO 2804; 428bp; English.
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
SQ Sequence 183 AA:
Query Match 69.2%; Score 18; DB 8; Length 183;
Best Local Similarity 20.0%; Pred. No. 3.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXXX 10
DB 36 CATTTC 45
RESULT 13
ABO76619
ID ABO76619 standard; protein; 191 AA.
XX
AC ABO76619;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #8794.
XX
KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
OS Pseudomonas aeruginosa.
XX
XX US6551795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
PA (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI WPI; 2003-615309/58.
DR N-PSDB; ABD10190.
XX
DR Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 25365; 455bp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 191 AA:
Query Match 69.2%; Score 18; DB 7; Length 191;
Best Local Similarity 20.0%; Pred. No. 3.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXXX 10
DB 49 CSTASATSC 58
RESULT 14
ADZ56213
ID ADZ56213 standard; protein; 214 AA.
XX
AC ADZ56213;
XX
DT 30-JUN-2005 (first entry)
XX
DE Human KIAA0779 splice variant clone C1AN00149041.a, protein.
XX
XX KIAA0779; gene expression; inflammation; antiinflammatory; cancer;
KM proliferation; neoplasm; cytoskeletal; immune disorder; immunomodulator;
KM metabolic disorder; metabolic; viral infection; virucide; infection.
XX
XX Homo sapiens.
XX
XX MO2005035569-A2.
XX
PD 21-APR-2005.
XX
XX 12-OCT-2004; 2004MO-US033408.
XX
XX 10-OCT-2003; 2003US-0510612P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Wong JGP, Hestir K, Colline ALT;
PI WPI; 2005-296268/30.
XX

DR N-PSDB; ADZ56200, ADZ56231.
XX
PT New isolated KIAA0779 nucleic acids and polypeptides, useful for
PT diagnosing, preventing and/or treating inflammatory, immune, viral
PT disorders and cancer, such as kidney, lung, ovarian, bladder, breast,
PT prostate and skin cancers.
XX
PS Claim 11; SEQ ID NO 18; 121pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule comprises at
CC least one polynucleotide sequence (appearing as ADZ56196-ADZ56197,
CC ADZ56199-ADZ56200, ADZ56226-ADZ56227 and ADZ56230-ADZ56231), sequences
CC hybridizing to them under high stringency conditions, sequences having at
CC least 80% sequence identity to them, their complements or biologically
CC active fragments. The nucleic acids are splice variants of the human
CC KIAA0779 gene, encoding transmembrane domain protein(s). Also included
CC are a double-stranded isolated nucleic acid molecule comprising the
CC nucleic acid molecule cited above, a vector comprising the isolated
CC expression of the nucleic acid molecule), a recombinant host cell
CC comprising the nucleic acid molecule cited above, an isolated polypeptide
CC (comprising the encoded proteins or the non-transmembrane (TM) regions of
CC the proteins), a method of making a recombinant host cell, a method of
CC making a polypeptide, a method of determining the presence of the nucleic
CC acid molecule cited above in a sample, a method of determining the
CC presence of a specific antibody to the polypeptide of (4) in a sample, a
CC method of determining the presence of the polypeptide of (4) in a sample, an
CC antibody specifically binding to and/or interfering with the biological
CC activity of the nucleic acid molecule cited above (or the polypeptide or
CC its biologically active fragment), a composition comprising a
CC pharmaceutical carrier or excipient (and one or more active agents chosen
CC from the nucleic acid molecule cited above, the vector, the polypeptide,
CC and the antibody), a bacteriophage comprising the antibody (or its
CC fragment), a bacterial cell comprising the bacteriophage, a recombinant
CC host cell that produces the antibody, an animal injected with one or more
CC active agents (chosen from the nucleic acid molecule, the vector, the
CC host cell, the polypeptide, and the antibody), a diagnostic kit
CC (comprising a nucleic acid molecule having at least 6 contiguous
CC nucleotides from the nucleic acid molecule cited above, the isolated
CC polypeptide, or antibody, and reagents to carry out an immunoassay), a
CC method of making an antibody, a method of identifying a modulating agent
CC that modulates the biological activity of the polypeptide, a modulator
CC composition comprising a modulator and a pharmaceutical carrier, a method
CC of treating a disease in a subject, and a method of treating cancer,
CC preferably kidney, cervical, squamous lung, ovarian, bladder, breast,
CC endometrial, prostate, and skin cancer, in a subject. The methods and
CC compositions of the present invention are useful for the diagnosis,
CC prevention and/or treatment of inflammatory, immune, viral disorders and
CC cancer, such as kidney, cervical, squamous lung, ovarian, bladder,
CC breast, endometrial, prostate, and skin cancer. The present sequence
CC represents a protein encoded by a cDNA splice variant from the KIAA0779
CC gene.
XX
SQ Sequence 214 AA;
XX
Query Match 69.2%; Score 18; DB 9; Length 214;
Best Local Similarity 20.0%; Pred. No. 3.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXC 10
Db 173 CAAAAAAC 182
RESULT 15
ADP30921
ID ADP30921 standard; protein; 228 AA.
XX
AC ADP30921;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1688.

XX
KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Viaricide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.

PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halebek RF, Huang MM, Kochakota S, Haisan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2919; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
XX Sequence 228 AA;
SQ
Query Match 69.2%; Score 18; DB 8; Length 228;
Best Local Similarity 20.0%; Pred. No. 3.8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXXX 10
DB 54 CATMAATTAC 63
RESULT 16
ADP31485
ID ADP31485 standard; protein; 252 AA.
XX
XX ADP31485;
AC
XX
XX 12-AUG-2004 (first entry)
DT
XX
XX Human secreted protein SEQ ID #2252.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
OS
XX
XX WO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PF
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0410963P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411033P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472430P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halebek RF, Huang MM, Kochakota S, Haisan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3483; 428bp; English.
XX

CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 252 AA:
Query Match 69.2%; Score 18; DB 8; Length 252;
Best Local Similarity 20.0%; Pred. No. 3.9;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXC 10
DB 210 CTTTATATAC 219
RESULT 17
ADP30479
ID ADP30479 standard; protein; 258 AA.
XX
AC ADP30479;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1246.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.

PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467209P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471366P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485252P.
PR 14-JUL-2003; 2003US-0486486P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PA Williams LT, Chu K, Lee B, Hestir K, Beaureang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2477; 428bp; English.
XX
PS
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 258 AA:
Query Match 69.2%; Score 18; DB 8; Length 258;
Best Local Similarity 20.0%; Pred. No. 3.9;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXC 10
DB 237 CTTTATATAC 246
RESULT 18
ADP31412

ID ADP31412 standard; protein; 264 AA.
 XX
 AC ADP31412;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human secreted protein SEQ ID #2179.
 XX
 KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
 cancer; inflammatory; immune; human secreted protein.
 OS Homo sapiens.
 XX
 PN WO2004035732-A2.
 XX
 PD 29-APR-2004.
 XX
 PF 28-AUG-2003; 2003WO-US026780.
 XX
 XX 29-AUG-2002; 2002US-0406576P.
 PR 29-AUG-2002; 2002US-0406579P.
 PR 29-AUG-2002; 2002US-0406585P.
 PR 29-AUG-2002; 2002US-0406588P.
 PR 29-AUG-2002; 2002US-0406608P.
 PR 29-AUG-2002; 2002US-0406611P.
 PR 29-AUG-2002; 2002US-0406612P.
 PR 29-AUG-2002; 2002US-0406616P.
 PR 29-AUG-2002; 2002US-0406640P.
 PR 29-AUG-2002; 2002US-0406642P.
 PR 29-AUG-2002; 2002US-0406646P.
 PR 29-AUG-2002; 2002US-0406653P.
 PR 29-AUG-2002; 2002US-0406655P.
 PR 29-AUG-2002; 2002US-0406666P.
 PR 17-SEP-2002; 2002US-0410946P.
 PR 17-SEP-2002; 2002US-0410947P.
 PR 17-SEP-2002; 2002US-0410948P.
 PR 17-SEP-2002; 2002US-0410949P.
 PR 17-SEP-2002; 2002US-0410953P.
 PR 17-SEP-2002; 2002US-0410953P.
 PR 17-SEP-2002; 2002US-0410957P.
 PR 17-SEP-2002; 2002US-0410958P.
 PR 17-SEP-2002; 2002US-0410959P.
 PR 17-SEP-2002; 2002US-0410960P.
 PR 17-SEP-2002; 2002US-0410961P.
 PR 17-SEP-2002; 2002US-0410962P.
 PR 17-SEP-2002; 2002US-0411019P.
 PR 17-SEP-2002; 2002US-0411022P.
 PR 17-SEP-2002; 2002US-0411023P.
 PR 17-SEP-2002; 2002US-0411024P.
 PR 17-SEP-2002; 2002US-0411032P.
 PR 17-SEP-2002; 2002US-0411035P.
 PR 17-SEP-2002; 2002US-0411037P.
 PR 17-SEP-2002; 2002US-0411041P.
 PR 17-SEP-2002; 2002US-0411045P.
 PR 17-SEP-2002; 2002US-0411046P.
 PR 17-SEP-2002; 2002US-0411048P.
 PR 17-SEP-2002; 2002US-0411052P.
 PR 17-SEP-2002; 2002US-0411055P.
 PR 17-SEP-2002; 2002US-0411073P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411101P.
 PR 17-SEP-2002; 2002US-0411111P.
 PR 18-APR-2003; 2003US-0463700P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 18-APR-2003; 2003US-0463733P.
 PR 18-APR-2003; 2003US-0463739P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 02-MAY-2003; 2003US-0467230P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 19-MAY-2003; 2003US-0471336P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.
 PR 09-JUN-2003; 2003US-0476641P.
 PR 08-JUL-2003; 2003US-0485218P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485224P.
 PR 08-JUL-2003; 2003US-0485325P.
 PR 14-JUL-2003; 2003US-0486446P.
 PR 14-JUL-2003; 2003US-0486480P.
 PR 15-JUL-2003; 2003US-0486893P.
 PR 15-JUL-2003; 2003US-0486960P.
 PR 08-AUG-2003; 2003US-0493341P.
 PR 08-AUG-2003; 2003US-0493370P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.
 XX
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
 XX
 PI Williams LT, Chu K, Lee E, Hestir K, Beaureang PA, Behrens D,
 PI Halenbeck RF, Huang MM, Kothakota S, Hsiehan L, Linnemann T;
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
 XX
 DR WPI; 2004-348438/32.
 XX
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 PT genetic, bacterial and viral diseases.
 XX
 PS Claim 1; SEQ ID NO 3410; 428bp; English.
 XX
 CC The present invention relates to an isolated nucleic acid molecule
 CC encoding a polypeptide which is believed to be cytostatic,
 CC antinflammatory, immunosuppressive, antibacterial and virucidal. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is
 CC available on WIPOMEB and is not in the specification.
 XX
 SQ Sequence 264 AA;
 Query Match 69.2%; Score 18; DB 8; Length 264;
 Best Local Similarity 20.0%; Pred. No. 3.9;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 CXXXXXXXC 10
 Db 186 CATTAATTC 195
 RESULT 19
 ADP30500
 ID ADP30500 standard; protein; 270 AA.
 XX
 AC ADP30500;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human secreted protein SEQ ID #1267.
 XX
 KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
 cancer; inflammatory; immune; human secreted protein.
 OS Homo sapiens.
 XX
 PN WO2004035732-A2.
 XX
 PD 29-APR-2004.
 XX
 PF 28-AUG-2003; 2003WO-US026780.
 XX
 XX 29-AUG-2002; 2002US-0406576P.
 PR 29-AUG-2002; 2002US-0406579P.
 PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486906P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Heslir K, Beaunrang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,
DR WFI; 2004-348438/32.

XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2498; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic.
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMED and is not in the specification.
XX
SQ Sequence 270 AA;
Query Match 69.2%; Score 18; DB 8; Length 270;
Best Local Similarity 20.0%; Pred. No. 3.9;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXC 10
Db 233 CTTATATATC 242
RESULT 20
ADP31321
ID ADP31321 standard; protein; 270 AA.
XX
AC ADP31321;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2088.
XX
DE Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN MO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485335P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI, 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1, SEQ ID NO 3319; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
XX Sequence 270 AA:

Query Match 69.2%; Score 18; DB 8; Length 270;
Best Local Similarity 20.0%; Pred. No. 3.9;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10
DB 245 CTTTATATAC 254

RESULT 21
ID ADP31473
ID ADP31473 standard; protein; 294 AA.
XX
AC ADP31473;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2240.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN MO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472420P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485355P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halsebeck RF, Huang MW, Kochakota S, Halsehan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3471; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
XX Sequence 294 AA;
SQ
Query Match 69.2%; Score 18; DB 8; Length 294;
Best Local Similarity 20.0%; Pred. No. 4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXXX 10
DB 72 CATTAAATC 81
RESULT 22
ADP31192
ID ADP31192 standard; protein; 297 AA.
XX
XX ADP31192;
AC
XX
XX 12-AUG-2004 (first entry)
DT
XX
XX Human secreted protein SEQ ID #1959.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX

PD 29-APR-2004.
XX
XX 28-AUG-2003; 2003MO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 29-AUG-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411054P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411101P.
XX 17-SEP-2002; 2002US-0411111P.
XX 18-APR-2003; 2003US-0463700P.
XX 18-APR-2003; 2003US-0463708P.
XX 18-APR-2003; 2003US-0463716P.
XX 18-APR-2003; 2003US-0463732P.
XX 02-MAY-2003; 2003US-0467199P.
XX 02-MAY-2003; 2003US-0467201P.
XX 02-MAY-2003; 2003US-0467203P.
XX 02-MAY-2003; 2003US-0467230P.
XX 19-MAY-2003; 2003US-0471306P.
XX 19-MAY-2003; 2003US-0471366P.
XX 22-MAY-2003; 2003US-0472420P.
XX 22-MAY-2003; 2003US-0472430P.
XX 09-JUN-2003; 2003US-0476609P.
XX 09-JUN-2003; 2003US-0476641P.
XX 08-JUL-2003; 2003US-0485218P.
XX 08-JUL-2003; 2003US-0485223P.
XX 08-JUL-2003; 2003US-0485224P.
XX 08-JUL-2003; 2003US-0485355P.
XX 14-JUL-2003; 2003US-0486446P.
XX 14-JUL-2003; 2003US-0486480P.
XX 15-JUL-2003; 2003US-0486891P.
XX 15-JUL-2003; 2003US-0486960P.
XX 08-AUG-2003; 2003US-0493341P.
XX 08-AUG-2003; 2003US-0493370P.
XX 08-AUG-2003; 2003US-0493573P.
XX 08-AUG-2003; 2003US-0493577P.
XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halembeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3190; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPower and is not in the specification.
XX
XX Sequence 297 AA;
SQ
Query Match 69.2%; Score 18; DB 8; Length 297;
Best Local Similarity 20.0%; Pred. No. 4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXC 10
DB 166 CTTATTATTC 175
DE Human protein sequence SEQ ID NO:800.
RESULT 23
AAM25285
ID AAM25285 standard; protein; 328 AA.
XX
XX AAM25285;
AC
XX
XX 16-OCT-2001 (first entry)
DT
XX
XX Human protein sequence SEQ ID NO:800.
DE
XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
XX antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
XX antibacterial; endocrine; cardiac; central nervous system; virucide;
XX anti-HIV; fungicide; antimutagen; cardiovascular; antiaemic; anaemia;
XX antiaggregant; haemostatic; vulnary; antileuk; osteopathic; eczema;
XX dermatological; antiallergic; antiaesthetic; antidiabetic; cytostatic;
XX neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
XX immunostimulant; gene therapy; antileuk therapy; vaccine; inflammation;
XX antiapophytic; rheumatoid arthritis; septic shock; pancreatitis;
XX cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
XX genetic disease; haematopoietic disorder; platelet disorder; asthma;
XX thrombocytopenia; osteoporosis; severe combined immunodeficiency;
XX allergic rhinitis; diabetes; multiple sclerosis; depression;
XX Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
XX neurological disorder.
XX
XX Homo sapiens.
OS
XX
XX WO200153455-A2.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 22-DEC-2000; 2000WO-US035017.
PF
XX
XX 23-DEC-1999; 99US-00471275.
PR
XX 21-JAN-2000; 2000US-0048725.
PR
XX 25-APR-2000; 2000US-00552317.
PR
XX
XX (HYSE-) HYSEQ INC.
PA

XX
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI; 2001-457603/49.
DR
XX N-PSDB; AAM99226.
DR
XX
XX Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX
XX Claim 20; Page 187; 1217bp; English.
XX
XX AAM99166 to AAM99904 encode the human proteins given in AAM25225 to
XX AAM25963. The proteins can have activities based on the tissues and cells
CC they are expressed in, such as: antiinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antiaemic; antiaggregant; haemostatic; vulnary;
CC antidiabetic; cytostatic; dermatological; antiallergic; antiaesthetic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antileuk therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders
XX
XX Sequence 328 AA;
SQ
Query Match 69.2%; Score 18; DB 4; Length 328;
Best Local Similarity 20.0%; Pred. No. 4.1;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXC 10
DB 63 CAAAAAAAC 72
DE
XX
XX ADP30702 standard; protein; 339 AA.
ID
XX
XX ADP30702;
AC
XX
XX 12-AUG-2004 (first entry)
DT
XX
XX Human secreted protein SEQ ID #1469.
DE
XX
XX Cytostatic; Antiinflammatory; immunosuppressive; antibacterial; virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
OS
XX
XX WO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PF
XX
XX 29-AUG-2002; 2002US-0406576P.
PR
XX 29-AUG-2002; 2002US-0406579P.
PR
XX 29-AUG-2002; 2002US-0406585P.
PR
XX 29-AUG-2002; 2002US-0406588P.
PR
XX 29-AUG-2002; 2002US-0406608P.
PR
XX 29-AUG-2002; 2002US-0406611P.
PR
XX 29-AUG-2002; 2002US-0406612P.
PR
XX 29-AUG-2002; 2002US-0406616P.
PR

PR 29-AUG-2002; 2002US-0406640P.
 PR 29-AUG-2002; 2002US-0406642P.
 PR 29-AUG-2002; 2002US-0406646P.
 PR 29-AUG-2002; 2002US-0406653P.
 PR 29-AUG-2002; 2002US-0406655P.
 PR 29-AUG-2002; 2002US-0406666P.
 PR 17-SEP-2002; 2002US-0410946P.
 PR 17-SEP-2002; 2002US-0410947P.
 PR 17-SEP-2002; 2002US-0410948P.
 PR 17-SEP-2002; 2002US-0410949P.
 PR 17-SEP-2002; 2002US-0410953P.
 PR 17-SEP-2002; 2002US-0410957P.
 PR 17-SEP-2002; 2002US-0410958P.
 PR 17-SEP-2002; 2002US-0410959P.
 PR 17-SEP-2002; 2002US-0410960P.
 PR 17-SEP-2002; 2002US-0410961P.
 PR 17-SEP-2002; 2002US-0410962P.
 PR 17-SEP-2002; 2002US-0411019P.
 PR 17-SEP-2002; 2002US-0411022P.
 PR 17-SEP-2002; 2002US-0411023P.
 PR 17-SEP-2002; 2002US-0411024P.
 PR 17-SEP-2002; 2002US-0411032P.
 PR 17-SEP-2002; 2002US-0411033P.
 PR 17-SEP-2002; 2002US-0411037P.
 PR 17-SEP-2002; 2002US-0411041P.
 PR 17-SEP-2002; 2002US-0411045P.
 PR 17-SEP-2002; 2002US-0411046P.
 PR 17-SEP-2002; 2002US-0411048P.
 PR 17-SEP-2002; 2002US-0411052P.
 PR 17-SEP-2002; 2002US-0411055P.
 PR 17-SEP-2002; 2002US-0411073P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411101P.
 PR 17-SEP-2002; 2002US-0411111P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463709P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 02-MAY-2003; 2003US-0467230P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 19-MAY-2003; 2003US-0471336P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 22-MAY-2003; 2003US-0472430P.
 PR 09-JUN-2003; 2003US-0476609P.
 PR 09-JUN-2003; 2003US-0476641P.
 PR 08-JUL-2003; 2003US-0485218P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485224P.
 PR 08-JUL-2003; 2003US-0485325P.
 PR 14-JUL-2003; 2003US-0486446P.
 PR 14-JUL-2003; 2003US-0486480P.
 PR 15-JUL-2003; 2003US-0486891P.
 PR 15-JUL-2003; 2003US-0486960P.
 PR 08-AUG-2003; 2003US-0493341P.
 PR 08-AUG-2003; 2003US-0493370P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.
 XX
 XX
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
 XX
 PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
 PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T,
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
 XX
 DR WPI; 2004-348438/32.
 XX
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 PT genetic, bacterial and viral diseases.
 XX

PS Claim 1; SEQ ID NO 2700; 428bp; English.
 XX
 CC The present invention relates to an isolated nucleic acid molecule
 CC encoding a polypeptide which is believed to be cytostatic,
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is
 CC available on WIPOMB and is not in the specification.
 XX
 SQ Sequence 339 AA;
 Query Match 69.2%; Score 18; DB 8; Length 339;
 Best Local Similarity 20.0%; Pred. No. 4.1;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 CXXXXXXXXX 10
 Db 89 CAATTAATTC 98
 RESULT 25
 ID ADP31441 standard; protein; 348 AA.
 XX
 AC ADP31441;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human secreted protein SEQ ID #2208.
 XX
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
 KW cancer; inflammatory; immune; human secreted protein.
 OS Homo sapiens.
 XX
 PN MO2004035732-A2.
 XX
 PD 29-APR-2004.
 XX
 XX 28-AUG-2003; 2003WO-US026780.
 XX
 PR 29-AUG-2002; 2002US-0406576P.
 PR 29-AUG-2002; 2002US-0406579P.
 PR 29-AUG-2002; 2002US-0406585P.
 PR 29-AUG-2002; 2002US-0406589P.
 PR 29-AUG-2002; 2002US-0406608P.
 PR 29-AUG-2002; 2002US-0406611P.
 PR 29-AUG-2002; 2002US-0406616P.
 PR 29-AUG-2002; 2002US-0406640P.
 PR 29-AUG-2002; 2002US-0406642P.
 PR 29-AUG-2002; 2002US-0406646P.
 PR 29-AUG-2002; 2002US-0406653P.
 PR 29-AUG-2002; 2002US-0406655P.
 PR 29-AUG-2002; 2002US-0406659P.
 PR 17-SEP-2002; 2002US-0410947P.
 PR 17-SEP-2002; 2002US-0410949P.
 PR 17-SEP-2002; 2002US-0410948P.
 PR 17-SEP-2002; 2002US-0410949P.
 PR 17-SEP-2002; 2002US-0410953P.
 PR 17-SEP-2002; 2002US-0410957P.
 PR 17-SEP-2002; 2002US-0410958P.
 PR 17-SEP-2002; 2002US-0410959P.
 PR 17-SEP-2002; 2002US-0410960P.
 PR 17-SEP-2002; 2002US-0410961P.
 PR 17-SEP-2002; 2002US-0410962P.
 PR 17-SEP-2002; 2002US-0411019P.
 PR 17-SEP-2002; 2002US-0411022P.
 PR 17-SEP-2002; 2002US-0411023P.
 PR 17-SEP-2002; 2002US-0411024P.
 PR 17-SEP-2002; 2002US-0411032P.

PR 17-SEP-2002; 2002US-0411035P.
 PR 17-SEP-2002; 2002US-0411037P.
 PR 17-SEP-2002; 2002US-0411041P.
 PR 17-SEP-2002; 2002US-0411045P.
 PR 17-SEP-2002; 2002US-0411046P.
 PR 17-SEP-2002; 2002US-0411048P.
 PR 17-SEP-2002; 2002US-0411052P.
 PR 17-SEP-2002; 2002US-0411055P.
 PR 17-SEP-2002; 2002US-0411073P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411101P.
 PR 17-SEP-2002; 2002US-0411111P.
 PR 18-APR-2003; 2003US-0463700P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 02-MAY-2003; 2003US-0467230P.
 PR 19-MAY-2003; 2003US-0471366P.
 PR 19-MAY-2003; 2003US-0471366P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 22-MAY-2003; 2003US-0472430P.
 PR 09-JUN-2003; 2003US-0476609P.
 PR 09-JUN-2003; 2003US-0476641P.
 PR 08-JUL-2003; 2003US-0485218P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485224P.
 PR 08-JUL-2003; 2003US-0485325P.
 PR 14-JUL-2003; 2003US-0486446P.
 PR 14-JUL-2003; 2003US-0486480P.
 PR 15-JUL-2003; 2003US-0486911P.
 PR 15-JUL-2003; 2003US-0486960P.
 PR 08-AUG-2003; 2003US-0493341P.
 PR 08-AUG-2003; 2003US-0493370P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
 XX
 PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
 PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;
 PI Pierre K, Wang Y, Wong JGP, Wu G, Zhang H;
 XX
 DR WPI; 2004-348438/32.
 XX
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 PT genetic, bacterial and viral diseases.
 XX
 PS Claim 1; SEQ ID NO 3439; 428bp; English.
 XX
 CC The present invention relates to an isolated nucleic acid molecule
 CC encoding a polypeptide which is believed to be cytostatic,
 CC antiinflammatory, immunosuppressive, antibacterial and antiviral. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is
 CC available on WIPOWEB and is not in the specification.
 XX
 SQ Sequence 348 AA;
 Query Match 69.2%; Score 18; DB 8; Length 348;
 Best Local Similarity 20.0%; Pred. No. 4.1;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

RESULT 26
 ABG21039
 ID ABG21039 standard; protein; 356 AA.
 XX
 AC ABG21039;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #21030.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW Food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS85226.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 51398; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences
 XX
 SQ Sequence 356 AA;
 Query Match 69.2%; Score 18; DB 4; Length 356;
 Best Local Similarity 20.0%; Pred. No. 4.2;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

RESULT 27
 ADP31267
 Db 183 CAABAAAAC 192

ID ADP31267 standard; protein; 357 AA.
 XX
 AC ADP31267;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human secreted protein SEQ ID #2034.
 XX
 KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
 cancer; inflammatory; immune; human secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO2004035732-A2.
 XX
 PD 29-APR-2004.
 XX
 PF 28-AUG-2003; 2003WO-US026780.
 XX
 PR 29-AUG-2002; 2002US-0406576P.
 PR 29-AUG-2002; 2002US-0406579P.
 PR 29-AUG-2002; 2002US-0406585P.
 PR 29-AUG-2002; 2002US-0406588P.
 PR 29-AUG-2002; 2002US-0406608P.
 PR 29-AUG-2002; 2002US-0406611P.
 PR 29-AUG-2002; 2002US-0406612P.
 PR 29-AUG-2002; 2002US-0406640P.
 PR 29-AUG-2002; 2002US-0406642P.
 PR 29-AUG-2002; 2002US-0406646P.
 PR 29-AUG-2002; 2002US-0406653P.
 PR 29-AUG-2002; 2002US-0406655P.
 PR 29-AUG-2002; 2002US-0406666P.
 PR 17-SEP-2002; 2002US-0410946P.
 PR 17-SEP-2002; 2002US-0410947P.
 PR 17-SEP-2002; 2002US-0410948P.
 PR 17-SEP-2002; 2002US-0410949P.
 PR 17-SEP-2002; 2002US-0410953P.
 PR 17-SEP-2002; 2002US-0410957P.
 PR 17-SEP-2002; 2002US-0410958P.
 PR 17-SEP-2002; 2002US-0410959P.
 PR 17-SEP-2002; 2002US-0410960P.
 PR 17-SEP-2002; 2002US-0410961P.
 PR 17-SEP-2002; 2002US-0410962P.
 PR 17-SEP-2002; 2002US-0411019P.
 PR 17-SEP-2002; 2002US-0411022P.
 PR 17-SEP-2002; 2002US-0411023P.
 PR 17-SEP-2002; 2002US-0411024P.
 PR 17-SEP-2002; 2002US-0411032P.
 PR 17-SEP-2002; 2002US-0411035P.
 PR 17-SEP-2002; 2002US-0411037P.
 PR 17-SEP-2002; 2002US-0411041P.
 PR 17-SEP-2002; 2002US-0411045P.
 PR 17-SEP-2002; 2002US-0411046P.
 PR 17-SEP-2002; 2002US-0411048P.
 PR 17-SEP-2002; 2002US-0411052P.
 PR 17-SEP-2002; 2002US-0411055P.
 PR 17-SEP-2002; 2002US-0411073P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411101P.
 PR 17-SEP-2002; 2002US-0411111P.
 PR 18-APR-2003; 2003US-0463700P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 02-MAY-2003; 2003US-0467230P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 19-MAY-2003; 2003US-0471366P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.
 PR 09-JUN-2003; 2003US-0476641P.
 PR 08-JUL-2003; 2003US-0485218P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485224P.
 PR 08-JUL-2003; 2003US-0485252P.
 PR 14-JUL-2003; 2003US-0486466P.
 PR 14-JUL-2003; 2003US-0486480P.
 PR 15-JUL-2003; 2003US-0486891P.
 PR 15-JUL-2003; 2003US-0486960P.
 PR 08-AUG-2003; 2003US-0493341P.
 PR 08-AUG-2003; 2003US-0493370P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.
 XX
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
 XX
 PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
 PI Halenbeck RF, Huang MM, Kothakota S, Halsehan L, Linnemann T;
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
 XX
 DR MPI; 2004-348438/32.
 XX
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 PT genetic, bacterial and viral diseases.
 XX
 PS Claim 1; SEQ ID NO 3265; 428bp; English.
 XX
 CC The present invention relates to an isolated nucleic acid molecule
 CC encoding a polypeptide which is believed to be cytostatic,
 CC antinflammatory, immunosuppressive, antibacterial and virucidal. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is
 CC available on WIPOMEB and is not in the specification.
 XX
 SQ Sequence 357 AA;
 Query Match 69.2%; Score 18; DB 8; Length 357;
 Best Local Similarity 20.0%; Pred. No. 4.2;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 CXXXXXXXXX 10
 DB 14 CATATATATAC 23
 ADP30505
 ID ADP30505 standard; protein; 357 AA.
 XX
 AC ADP30505;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human secreted protein SEQ ID #1272.
 XX
 KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
 cancer; inflammatory; immune; human secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO2004035732-A2.
 XX
 PD 29-APR-2004.
 XX
 PF 28-AUG-2003; 2003WO-US026780.
 XX
 PR 29-AUG-2002; 2002US-0406576P.
 PR 29-AUG-2002; 2002US-0406579P.
 PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-048660P.
PR 15-JUL-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang WM, Kotlakota S, Haishan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,
XX
XX WPI; 2004-348438/32.

XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 2503; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMER and is not in the specification.
XX
SQ Sequence 357 AA;
XX
Query Match 69.2%; Score 18; DB 8; Length 357;
Best Local Similarity 20.0%; Pred. No. 4.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXC 10
Db 169 CTAATTTC 178
RESULT 29
ADP31439
ID ADP31439 standard; protein; 360 AA.
XX
AC ADP31439;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2206.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411011P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 02-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485225P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

XX DA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
XX PI Halenbeck RF, Huang MM, Kotchkota S, Haishan L, Linemann T;
XX PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,
XX WPI; 2004-348438/32.

XX DR New nucleic acid molecule for diagnosing, preventing or treating diseases
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX PT genetic, bacterial and viral diseases.

XX PS Claim 1; SEQ ID NO 3437; 428bp; English.

XX CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available from WIPOWEB and is not in the specification.

XX Sequence 360 AA;

QY	1	XXXXXXXXXX	10
Db	70	CTAATTAAAC	79
RESULT 30			
ID	ADP31218		
AC	ADP31218 standard; protein; 390 AA.		
XX			
XX	ADP31218;		
DT	12-AUG-2004	(first entry)	
XX			
DE	Human secreted protein SEQ ID #1985.		
XX			
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;		
KW	cancer; inflammatory; immune; human secreted protein.		
XX			
OS	Homo sapiens.		
XX			
PN	WO2004035732-A2.		
XX			
PD	29-APR-2004.		
PF	28-AUG-2003; 2003WO-US026780.		
XX			
PR	29-AUG-2002; 2002US-0406576P.		
PR	29-AUG-2002; 2002US-0406579P.		
PR	29-AUG-2002; 2002US-0406585P.		
PR	29-AUG-2002; 2002US-0406588P.		
PR	29-AUG-2002; 2002US-0406589P.		
PR	29-AUG-2002; 2002US-0406608P.		
PR	29-AUG-2002; 2002US-0406611P.		
PR	29-AUG-2002; 2002US-0406612P.		
PR	29-AUG-2002; 2002US-0406616P.		
PR	29-AUG-2002; 2002US-0406640P.		
PR	29-AUG-2002; 2002US-0406642P.		
PR	29-AUG-2002; 2002US-0406646P.		
PR	29-AUG-2002; 2002US-0406653P.		
PR	29-AUG-2002; 2002US-0406655P.		
PR	29-AUG-2002; 2002US-0406666P.		
PR	17-SEP-2002; 2002US-0410946P.		
PR	17-SEP-2002; 2002US-0410947P.		
PR	17-SEP-2002; 2002US-0410948P.		
PR	17-SEP-2002; 2002US-0410949P.		
PR	17-SEP-2002; 2002US-0410953P.		
PR	17-SEP-2002; 2002US-0410957P.		
PR	17-SEP-2002; 2002US-0410958P.		
PR	17-SEP-2002; 2002US-0410959P.		
PR	17-SEP-2002; 2002US-0410960P.		
PR	17-SEP-2002; 2002US-0410961P.		
PR	17-SEP-2002; 2002US-0410962P.		
PR	17-SEP-2002; 2002US-0411019P.		
PR	17-SEP-2002; 2002US-0411022P.		
PR	17-SEP-2002; 2002US-0411023P.		
PR	17-SEP-2002; 2002US-0411024P.		
PR	17-SEP-2002; 2002US-0411032P.		
PR	17-SEP-2002; 2002US-0411037P.		
PR	17-SEP-2002; 2002US-0411037P.		
PR	17-SEP-2002; 2002US-0411041P.		
PR	17-SEP-2002; 2002US-0411045P.		
PR	17-SEP-2002; 2002US-0411046P.		
PR	17-SEP-2002; 2002US-0411048P.		
PR	17-SEP-2002; 2002US-0411052P.		
PR	17-SEP-2002; 2002US-0411055P.		
PR	17-SEP-2002; 2002US-0411073P.		
PR	17-SEP-2002; 2002US-0411082P.		
PR	17-SEP-2002; 2002US-0411101P.		
PR	17-SEP-2002; 2002US-0411111P.		
PR	18-APR-2003; 2003US-0463700P.		
PR	18-APR-2003; 2003US-0463708P.		
PR	18-APR-2003; 2003US-0463716P.		
PR	18-APR-2003; 2003US-0463732P.		
PR	02-MAY-2003; 2003US-0467199P.		

PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471306P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RT, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3216; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antitumorigenic, immunosuppressive, antibacterial and antiviral. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 390 AA:
Query Match 69.2%; Score 18; DB 8; Length 390;
Best Local Similarity 20.0%; Pred. No. 4.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXXX 10
Db 60 CAAAAAATTC 69
RESULT 31
ID ADP31159 standard; protein; 421 AA.
XX
XX ADP31159;
XX
XX 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #1926.
XX
XX Cytostatic; Antitumorigenic; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX OS
XX PN WO2004035732-A2.
XX

PD 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411101P.
XX 17-SEP-2002; 2002US-0411111P.
XX 18-APR-2003; 2003US-0463700P.
XX 18-APR-2003; 2003US-0463708P.
XX 18-APR-2003; 2003US-0463716P.
XX 18-APR-2003; 2003US-0463732P.
XX 02-MAY-2003; 2003US-0467199P.
XX 02-MAY-2003; 2003US-0467201P.
XX 02-MAY-2003; 2003US-0467203P.
XX 02-MAY-2003; 2003US-0467230P.
XX 02-MAY-2003; 2003US-0471306P.
XX 19-MAY-2003; 2003US-0471366P.
XX 19-MAY-2003; 2003US-0472430P.
XX 22-MAY-2003; 2003US-0472430P.
XX 09-JUN-2003; 2003US-0476609P.
XX 09-JUN-2003; 2003US-0476641P.
XX 08-JUL-2003; 2003US-0485218P.
XX 08-JUL-2003; 2003US-0485223P.
XX 08-JUL-2003; 2003US-0485224P.
XX 08-JUL-2003; 2003US-0485225P.
XX 14-JUL-2003; 2003US-0486446P.
XX 14-JUL-2003; 2003US-0486480P.
XX 15-JUL-2003; 2003US-0486891P.
XX 15-JUL-2003; 2003US-0486960P.
XX 08-AUG-2003; 2003US-0493341P.
XX 08-AUG-2003; 2003US-0493370P.
XX 08-AUG-2003; 2003US-0493573P.
XX 08-AUG-2003; 2003US-0493577P.
XX

Query Match	Best Local Similarity	Score 18;	DB 8;	Length 421;
Matches 2;	Conservative	0;	Mismatches 8;	Indels 0;
Gaps	0;			
1	CXXXXXXXXC	10		
385	CAAAAATC	394		
RESULT 32				
ADP31495	ADP31495	standard; protein; 426 AA.		
ADP31495;				
12-AUG-2004	(first entry)			
Human secreted protein SEQ ID #2262.				
Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide; cancer; inflammatory; immune; human secreted protein.				
Homo sapiens.				
WO2004035732-A2.				
29-APR-2004.				
28-AUG-2003;	2003WO-US026780.			
29-AUG-2002;	2002US-0406576P.			
29-AUG-2002;	2002US-0406579P.			
29-AUG-2002;	2002US-0406585P.			
29-AUG-2002;	2002US-0406588P.			
29-AUG-2002;	2002US-0406608P.			
29-AUG-2002;	2002US-0406611P.			
29-AUG-2002;	2002US-0406612P.			
29-AUG-2002;	2002US-0406616P.			
29-AUG-2002;	2002US-0406640P.			
29-AUG-2002;	2002US-0406642P.			
29-AUG-2002;	2002US-0406646P.			
29-AUG-2002;	2002US-0406653P.			
29-AUG-2002;	2002US-0406655P.			
29-AUG-2002;	2002US-0406666P.			
17-SEP-2002;	2002US-0410946P.			
17-SEP-2002;	2002US-0410947P.			
17-SEP-2002;	2002US-0410948P.			
17-SEP-2002;	2002US-0410949P.			

PR 17-SEP-2002; 2002US-0410953B.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471306P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476809P.
PR 09-JUN-2003; 2003US-0476841P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RP, Huang MM, Kotnakota S, Haislan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
DR
XX
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3493; 428pp; English.

The present invention relates to an isolated nucleic acid molecule encoding a polypeptide which is believed to be cytosolic. The anti-inflammatory, immunosuppressive, antibacterial and virucidal. The composition and methods are useful for diagnosing, preventing and treating diseases such as proliferative (e.g. cancer), inflammatory, immune, metabolic, genetic, bacterial and viral diseases. The present sequence represents a human secreted protein. The present sequence is available on WIPWEB and is not in the specification.

XX SQ Sequence 426 AA;
Query Match 69.2%; Score 18; DB 8; Length 426;
Best Local Similarity 20.0%; Pred. No. 4.3;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXXX 10
DB 66 CATTAAATTC 75
RESULT 33
ABO82561
ID ABO82561 standard; protein; 438 AA.
XX AC ABO82561;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polypeptide #14736.
XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX DR WPI: 2003-615309/58.
XX DR N-PSDB; ABD16132.
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.
XX PS Disclosure; SEQ ID NO 31307; 455pp; English.
XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the
XX CC polynucleotides encoding them. The sequences are useful in diagnosis and
XX CC therapy of pathological conditions, as molecular targets for diagnostics,
XX CC prophylaxis and treatment of pathological conditions resulting from a
XX CC bacterial infection, for evaluating a compound, such as a polypeptide,
XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of
XX CC effective antibacterial targets, as targets for antibacterial drugs,
XX CC including anti-P. aeruginosa drugs, as templates for recombinant
XX CC production of P. aeruginosa-derived peptides or polypeptides, as target
XX CC components for diagnosis and/or treatment of P. aeruginosa-caused
XX CC infection, and in detection of P. aeruginosa sequences or other sequences
XX CC of Pseudomonas species using biotech technology. Sequences ABO67826-
XX CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format from USPTO at
XX CC seqdata.uspto.gov/sequence.html
XX SQ Sequence 438 AA;
Query Match 69.2%; Score 18; DB 7; Length 438;
Best Local Similarity 20.0%; Pred. No. 4.3;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXXX 10

DB 56 CSAATASSTC 65
RESULT 34
ADP31465
ID ADP31465 standard; protein; 453 AA.
XX AC ADP31465;
XX DT 12-AUG-2004 (first entry)
XX DE Human secreted protein SEQ ID #2232.
XX DE Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX KW cancer; inflammatory; immune; human secreted protein.
XX OS Homo sapiens.
XX PN WO2004035732-A2.
XX PD 29-APR-2004.
XX PF 28-AUG-2003; 2003WO-US026780.
XX PR 29-AUG-2002; 2002US-0406576P.
XX PR 29-AUG-2002; 2002US-0406579P.
XX PR 29-AUG-2002; 2002US-0406585P.
XX PR 29-AUG-2002; 2002US-0406588P.
XX PR 29-AUG-2002; 2002US-0406608P.
XX PR 29-AUG-2002; 2002US-0406611P.
XX PR 29-AUG-2002; 2002US-0406612P.
XX PR 29-AUG-2002; 2002US-0406616P.
XX PR 29-AUG-2002; 2002US-0406640P.
XX PR 29-AUG-2002; 2002US-0406642P.
XX PR 29-AUG-2002; 2002US-0406646P.
XX PR 29-AUG-2002; 2002US-0406653P.
XX PR 29-AUG-2002; 2002US-0406655P.
XX PR 29-AUG-2002; 2002US-0410946P.
XX PR 17-SEP-2002; 2002US-0410947P.
XX PR 17-SEP-2002; 2002US-0410948P.
XX PR 17-SEP-2002; 2002US-0410949P.
XX PR 17-SEP-2002; 2002US-0410953P.
XX PR 17-SEP-2002; 2002US-0410957P.
XX PR 17-SEP-2002; 2002US-0410958P.
XX PR 17-SEP-2002; 2002US-0410959P.
XX PR 17-SEP-2002; 2002US-0410960P.
XX PR 17-SEP-2002; 2002US-0410961P.
XX PR 17-SEP-2002; 2002US-0410962P.
XX PR 17-SEP-2002; 2002US-0411019P.
XX PR 17-SEP-2002; 2002US-0411022P.
XX PR 17-SEP-2002; 2002US-0411024P.
XX PR 17-SEP-2002; 2002US-0411032P.
XX PR 17-SEP-2002; 2002US-0411035P.
XX PR 17-SEP-2002; 2002US-0411037P.
XX PR 17-SEP-2002; 2002US-0411041P.
XX PR 17-SEP-2002; 2002US-0411045P.
XX PR 17-SEP-2002; 2002US-0411046P.
XX PR 17-SEP-2002; 2002US-0411048P.
XX PR 17-SEP-2002; 2002US-0411052P.
XX PR 17-SEP-2002; 2002US-0411055P.
XX PR 17-SEP-2002; 2002US-0411073P.
XX PR 17-SEP-2002; 2002US-0411082P.
XX PR 17-SEP-2002; 2002US-0411101P.
XX PR 17-SEP-2002; 2002US-0411111P.
XX PR 18-APR-2003; 2003US-0463700P.
XX PR 18-APR-2003; 2003US-0463708P.
XX PR 18-APR-2003; 2003US-0463716P.
XX PR 18-APR-2003; 2003US-0463732P.
XX PR 02-MAY-2003; 2003US-0467199P.
XX PR 02-MAY-2003; 2003US-0467201P.
XX PR 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang WM, Kochakota S, Haislan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,
XX
DR MPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3463; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 453 AA:
Query Match 69.2%; Score 18; DB 8; Length 453;
Best Local Similarity 20.0%; Pred. No. 4.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXC 10
DB 312 CAATTAATC 321
RESULT 35
ADP30854
ID ADP30854 standard; protein; 471 AA.
XX
AC ADP30854;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1621.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN MO2004035732-A2.
XX
PD 29-APR-2004.
XX

PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411025P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411033P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX


```

Query Match          69.2%; Score 18; DB 8; Length 525;
Best Local Similarity 20.0%; Pred. NO. 4.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      1 CXXXXXXXXXC 10
        |             |
        475 CAAATTTTAC 484

Db

RESULT 37
ADP30855
ID      ADP30855 standard; protein; 549 AA.
AC      ADP30855;
XX
XX      12-AUG-2004 (first entry)
XX
XX      Human secreted protein SEQ ID #1622.
XX
XX      Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX      cancer; inflammatory; immune; human secreted protein.
XX
XX      Homo sapiens.
XX
XX      MO2004035732-A2.
XX
XX      29-APR-2004.
XX
XX      28-AUG-2003; 2003WO-US026780.
XX
XX      29-AUG-2002; 2002US-0406576P.
XX      29-AUG-2002; 2002US-0406579P.
XX      29-AUG-2002; 2002US-0406585P.
XX      29-AUG-2002; 2002US-0406588P.
XX      29-AUG-2002; 2002US-0406608P.
XX      29-AUG-2002; 2002US-0406611P.
XX      29-AUG-2002; 2002US-0406612P.
XX      29-AUG-2002; 2002US-0406616P.
XX      29-AUG-2002; 2002US-0406640P.
XX      29-AUG-2002; 2002US-0406642P.
XX      29-AUG-2002; 2002US-0406646P.
XX      29-AUG-2002; 2002US-0406653P.
XX      29-AUG-2002; 2002US-0406655P.
XX      29-AUG-2002; 2002US-0406666P.
XX      17-SEP-2002; 2002US-0410946P.
XX      17-SEP-2002; 2002US-0410947P.
XX      17-SEP-2002; 2002US-0410948P.
XX      17-SEP-2002; 2002US-0410949P.
XX      17-SEP-2002; 2002US-0410951P.
XX      17-SEP-2002; 2002US-0410953P.
XX      17-SEP-2002; 2002US-0410957P.
XX      17-SEP-2002; 2002US-0410958P.
XX      17-SEP-2002; 2002US-0410959P.
XX      17-SEP-2002; 2002US-0410960P.
XX      17-SEP-2002; 2002US-0410961P.
XX      17-SEP-2002; 2002US-0410962P.
XX      17-SEP-2002; 2002US-0411019P.
XX      17-SEP-2002; 2002US-0411022P.
XX      17-SEP-2002; 2002US-0411023P.
XX      17-SEP-2002; 2002US-0411024P.
XX      17-SEP-2002; 2002US-0411025P.
XX      17-SEP-2002; 2002US-0411035P.
XX      17-SEP-2002; 2002US-0411037P.
XX      17-SEP-2002; 2002US-0411041P.
XX      17-SEP-2002; 2002US-0411045P.
XX      17-SEP-2002; 2002US-0411046P.
XX      17-SEP-2002; 2002US-0411048P.
XX      17-SEP-2002; 2002US-0411052P.
XX      17-SEP-2002; 2002US-0411055P.
XX      17-SEP-2002; 2002US-0411073P.
XX      17-SEP-2002; 2002US-0411082P.
XX      17-SEP-2002; 2002US-0411101P.
XX      17-SEP-2002; 2002US-0411111P.

```

```

PR      18-APR-2003; 2003US-0463700P.
PR      18-APR-2003; 2003US-0463708P.
PR      18-APR-2003; 2003US-0463716P.
PR      18-APR-2003; 2003US-0463732P.
PR      02-MAY-2003; 2003US-0467199P.
PR      02-MAY-2003; 2003US-0467201P.
PR      02-MAY-2003; 2003US-0467203P.
PR      02-MAY-2003; 2003US-0467230P.
PR      19-MAY-2003; 2003US-0471306P.
PR      19-MAY-2003; 2003US-0471336P.
PR      22-MAY-2003; 2003US-0472420P.
PR      22-MAY-2003; 2003US-0472430P.
PR      09-JUN-2003; 2003US-0476609P.
PR      09-JUN-2003; 2003US-0476641P.
PR      08-JUL-2003; 2003US-0485218P.
PR      08-JUL-2003; 2003US-0485223P.
PR      08-JUL-2003; 2003US-0485224P.
PR      08-JUL-2003; 2003US-0485325P.
PR      14-JUL-2003; 2003US-048646P.
PR      14-JUL-2003; 2003US-0486480P.
PR      15-JUL-2003; 2003US-0486891P.
PR      15-JUL-2003; 2003US-0486960P.
PR      08-AUG-2003; 2003US-0493341P.
PR      08-AUG-2003; 2003US-0493370P.
PR      08-AUG-2003; 2003US-0493573P.
PR      08-AUG-2003; 2003US-0493577P.
XX
XX      (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX      Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
XX      Halenbeck RF, Huang MM, Kothakota S, Halshan L, Harnemann T;
XX      Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX      WPI; 2004-348438/32.
XX
XX      New nucleic acid molecule for diagnosing, preventing or treating diseases
XX      such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX      genetic, bacterial and viral diseases.
XX
XX      Claim 1; SEQ ID NO 2853; 428pp; English.
XX
XX      The present invention relates to an isolated nucleic acid molecule
XX      encoding a polypeptide which is believed to be cytostatic,
XX      antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX      composition and methods are useful for diagnosing, preventing and
XX      treating diseases such as proliferative (e.g. cancer), inflammatory,
XX      immune, metabolic, genetic, bacterial and viral diseases. The present
XX      sequence represents a human secreted protein. The present sequence is
XX      available on WIPOMEB and is not in the specification.
XX
XX      Sequence 549 AA;

Query Match          69.2%; Score 18; DB 8; Length 549;
Best Local Similarity 20.0%; Pred. NO. 4.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      1 CXXXXXXXXXC 10
        |             |
        220 CAAATTTATC 229

Db

RESULT 38
ADP31416
ID      ADP31416 standard; protein; 555 AA.
AC      ADP31416;
XX
XX      12-AUG-2004 (first entry)
XX
XX      Human secreted protein SEQ ID #2183.
XX
XX      Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX      cancer; inflammatory; immune; human secreted protein.

```

XX Homo sapiens.
OS
XX MO2004035732-A2.
PN
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406644P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463733P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.

PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang WM, Kothakota S, Halsehan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 3414; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytosratic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.
XX
SQ Sequence 555 AA;
XX
Query Match 69.2%; Score 18; DB 8; Length 555;
Best Local Similarity 20.0%; Pred. No. 4.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 1 CXXXXXXXXC 10
Db 211 CTTATTATTAC 220
RESULT 39
ADP31417
ID ADP31417 standard; protein; 555 AA.
XX
XX AC ADP31417;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2184.
XX
KW Cytosratic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
PN MO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411083P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463733P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486464P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Hwang MM, Kotchakota S, Hsieh L, Linnemann T,
PI Pierce K, Wang Y, Wong JGF, Wu G, Zhang H;
XX
XX WPT, 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PI genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3415; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPower and is not in the specification.
XX

XX Sequence 555 AA;

Query Match 69.2%; Score 18; DB 8; Length 555;
Best Local Similarity 20.0%; Pred. No. 4.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXX 10
Db 211 CTTTATTAC 220

RESULT 40
ADP31699
ID ADP31699 standard; protein; 588 AA.

AC ADP31699;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2466.

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

PN WO2004035732-A2.

PD 29-APR-2004.

PF 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.

PR	17-SEP-2002;	2002US-0411048P.
PR	17-SEP-2002;	2002US-0411052P.
PR	17-SEP-2002;	2002US-0411055P.
PR	17-SEP-2002;	2002US-0411073P.
PR	17-SEP-2002;	2002US-0411082P.
PR	17-SEP-2002;	2002US-0411101P.
PR	17-SEP-2002;	2002US-0411111P.
PR	18-APR-2003;	2003US-0463700P.
PR	18-APR-2003;	2003US-0463708P.
PR	18-APR-2003;	2003US-0463716P.
PR	18-APR-2003;	2003US-0463732P.
PR	02-MAY-2003;	2003US-0467199P.
PR	02-MAY-2003;	2003US-0467201P.
PR	02-MAY-2003;	2003US-0467203P.
PR	02-MAY-2003;	2003US-0467230P.
PR	19-MAY-2003;	2003US-0471306P.
PR	19-MAY-2003;	2003US-0471336P.
PR	22-MAY-2003;	2003US-0472420P.
PR	22-MAY-2003;	2003US-0472430P.
PR	09-JUN-2003;	2003US-0476609P.
PR	09-JUN-2003;	2003US-0476641P.
PR	08-JUL-2003;	2003US-0485218P.
PR	08-JUL-2003;	2003US-0485225P.
PR	08-JUL-2003;	2003US-0485224P.
PR	08-JUL-2003;	2003US-0485325P.
PR	14-JUL-2003;	2003US-0486446P.
PR	14-JUL-2003;	2003US-0486480P.
PR	15-JUL-2003;	2003US-0486891P.
PR	15-JUL-2003;	2003US-0486960P.
PR	08-AUG-2003;	2003US-0493341P.
PR	08-AUG-2003;	2003US-0493370P.
PR	08-AUG-2003;	2003US-0493573P.
PR	08-AUG-2003;	2003US-0493577P.

PI Williams JT, Chu K, Lee E, Hestir K, Beunang PA, Behrens D,
PI Haladenbeck RF, Huang MM, Kotchakota S, Hsieh L, Linnemann T;
PI Pierce Y, Wang JF, Wong JCF, Wu G, Zhang H;
XX WPI; 2004-348438/32.
DR
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 3697; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC anti-inflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.

Query Match	69.2%	Score 18;	DB 8;	Length 588;
Best Local Similarity	20.0%	Pred. No. 4.6;		
Matches 2; Conservative	0;	Mismatches 8;	Indels 0;	Gaps 0

Qy	1	XXXXXXXXXX	10
Db	277	CATTTTTTC	286

RESULT 41
ADZ56209
ID ADZ56209 standard; protein; 591 AA.
XX
AC ADZ56209;

30-JUN-2005 (first entry)

Human KIA00779 splice variant clone 24980850:24980849a, nonTM_1 #1.

KIA00779; gene expression; inflammation; antiinflammatory; cancer; proliferation; neoplasia; cytostatic; immune disorder; immunomodulator; metabolic disorder; metabolic; viral infection; virucide; infection.

Homo sapiens.

WO2005035569-A2.

21-APR-2005.

12-OCT-2004; 2004MO-US033408.

10-OCT-2003; 2003US-0510612P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

Wong JGP, Hestir K, Collins ALT;

WPI; 2005-296268/30.

New isolated KIA00779 nucleic acids and polypeptides, useful for diagnosing, preventing and/or treating inflammatory, immune, viral disorders and cancer, such as kidney, lung, ovarian, bladder, breast, prostate and skin cancers.

Claim 11; SEQ ID NO 14; 121pp; English.

The invention relates to an isolated nucleic acid molecule comprising at least one polynucleotide sequence (appearing as AD256196-AD256197, CC AD256199-AD256200, AD256226-AD256227 and AD256230-AD256231), sequences hybridizing to them under high stringency conditions, sequences having at least 80% sequence identity to them, their complements or biologically active fragments. The nucleic acids are splice variants of the human KIA00779 gene, encoding transmembrane domain protein(s). Also included are a double-stranded isolated nucleic acid molecule comprising the nucleic acid molecule cited above, a vector comprising the isolated nucleic acid molecule cited above, and a promoter that regulates the expression of the nucleic acid molecule), a recombinant host cell comprising the nucleic acid molecule cited above, an isolated polypeptide (comprising the encoded proteins or the non-transmembrane (TM) regions of the proteins), a method of making a recombinant host cell, a method of making a polypeptide, a method of determining the presence of the nucleic acid molecule cited above in a sample, a method of determining the presence of a specific antibody to the polypeptide of (4) in a sample, a method of determining the presence of the polypeptide in a sample, an antibody specifically binding to and/or interfering with the biological activity of the nucleic acid molecule cited above for the polypeptide or its biologically active fragment), a composition comprising a pharmaceutical carrier or excipient (and one or more active agents chosen from the nucleic acid molecule cited above, the vector, the polypeptide, and the antibody), a bacteriophage comprising the bacteriophage, a recombinant host cell that produces the antibody, an animal injected with one or more active agents (chosen from the nucleic acid molecule, the vector, the host cell, the polypeptide, and the antibody), a diagnostic kit (comprising a nucleic acid molecule having at least 6 contiguous nucleotides from the nucleic acid molecule cited above, the isolated polypeptide, or antibody, and reagents to carry out an immunoassay), a method of making an antibody, a method of identifying a modulating agent that modulates the biological activity of the polypeptide, a modulator composition comprising a modulator and a pharmaceutical carrier, a method of treating a disease in a subject, and a method of treating cancer, preferably kidney, cervical, squamous lung, ovarian, bladder, breast, endometrial, prostate, and skin cancer, in a subject. The methods and compositions of the present invention are useful for the diagnosis, prevention and/or treatment of inflammatory, immune, viral disorders and cancer, such as kidney, cervical, squamous lung, ovarian, bladder, breast, endometrial, prostate, and skin cancer. The present sequence

CC represents a non TM region of a protein encoded by a cDNA splice variant
 CC from the KIAA0779 gene.
 XX
 SQ Sequence 591 AA;

Query Match 69.2%; Score 18; DB 9; Length 591;
 Best Local Similarity 20.0%; Pred. No. 4.6;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 10
 DB 173 CAAAAAAAC 182

RESULT 42
 AD256223
 ID AD256223 standard; protein; 591 AA.

AC AD256223;
 DT 30-JUN-2005 (first entry)

DE Human KIAA0779 splice variant clone 24980850:249808498, nontm_1 #2.

KW KIAA0779; gene expression; inflammation; antiinflammatory; cancer;
 KW proliferation; neoplasm; cytoskeletal; immune disorder; immunomodulator;
 KW metabolic disorder; metabolic; viral infection; virucide; infection.

OS Homo sapiens.

PN WO2005035569-A2.

PD 21-APR-2005.

PE 12-OCT-2004; 2004WO-US033408.

PR 10-OCT-2003; 2003US-0510612P.

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Wong JGP, Heetir K, Collins ALT;

DR WPI; 2005-296268/30.

PT New isolated KIAA0779 nucleic acids and polypeptides, useful for
 PT diagnosing, preventing and/or treating inflammatory, immune, viral
 PT disorders and cancer, such as kidney, lung, ovarian, bladder, breast,
 PT prostate and skin cancers.

PS Claim 11; SEQ ID NO 28; 121bp; English.

CC The invention relates to an isolated nucleic acid molecule comprising at
 CC least one polynucleotide sequence (appearing as AD256196-AD256197,
 CC AD256199-AD256200, AD256226-AD256227 and AD256230-AD256231), sequences
 CC hybridizing to them under high stringency conditions, sequences having at
 CC least 80% sequence identity to them, their complements or biologically
 CC active fragments. The nucleic acids are splice variants of the human
 CC KIAA0779 gene, encoding transmembrane domain protein(s). Also included
 CC are a double-stranded isolated nucleic acid molecule comprising the
 CC nucleic acid molecule cited above, a vector comprising the isolated
 CC nucleic acid molecule cited above, (and a promoter that regulates the
 CC expression of the nucleic acid molecule), a recombinant host cell
 CC comprising the nucleic acid molecule cited above, an isolated polypeptide
 CC (comprising the encoded proteins or the non-transmembrane (TM) regions of
 CC the proteins), a method of making a recombinant host cell, a method of
 CC making a polypeptide, a method of determining the presence of the nucleic
 CC acid molecule cited above in a sample, a method of determining the
 CC presence of a specific antibody to the polypeptide of (4) in a sample, a
 CC method of determining the presence of the polypeptide in a sample, an
 CC antibody specifically binding to and/or interfering with the biological
 CC activity of the nucleic acid molecule cited above (or the polypeptide or
 CC its biologically active fragment), a composition comprising a
 CC pharmaceutical carrier or excipient (and one or more active agents chosen

CC from the nucleic acid molecule cited above, the vector, the polypeptide,
 CC and the antibody), a bacteriophage comprising the antibody (or its
 CC fragment), a bacterial cell comprising the bacteriophage, a recombinant
 CC host cell that produces the antibody, an animal injected with one or more
 CC active agents (chosen from the nucleic acid molecule, the vector, the
 CC host cell, the polypeptide, and the antibody), a diagnostic kit
 CC (comprising a nucleic acid molecule having at least 6 contiguous
 CC nucleotides from the nucleic acid molecule cited above, the isolated
 CC polypeptide, or antibody, and reagents to carry out an immunoassay), a
 CC method of making an antibody, a method of identifying a modulating agent
 CC that modulates the biological activity of the polypeptide, a modulator
 CC composition comprising a modulator and a pharmaceutical carrier, a method
 CC of treating a disease in a subject, and a method of treating cancer,
 CC preferably kidney, cervical, squamous lung, ovarian, bladder, breast,
 CC endometrial, prostate, and skin cancer, in a subject. The methods and
 CC compositions of the present invention are useful for the diagnosis,
 CC prevention and/or treatment of inflammatory, immune, viral disorders and
 CC cancer, such as kidney, cervical, squamous lung, ovarian, bladder,
 CC breast, endometrial, prostate, and skin cancer. The present sequence
 CC represents a non TM region of a protein encoded by a cDNA splice variant
 CC from the KIAA0779 gene.

SQ Sequence 591 AA;

Query Match 69.2%; Score 18; DB 9; Length 591;
 Best Local Similarity 20.0%; Pred. No. 4.6;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 10
 DB 173 CAAAAAAAC 182

RESULT 43

ADP30865
 ID ADP30865 standard; protein; 600 AA.

AC ADP30865;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #1632.

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
 KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

PN WO2004035732-A2.

PD 29-APR-2004.

PE 28-AUG-2003; 2003WO-US026780.

PR 29-AUG-2002; 2002US-0406576P.

PR 29-AUG-2002; 2002US-0406579P.

PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.

PR 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406616P.

PR 29-AUG-2002; 2002US-0406640P.

PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.

PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.

PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.

XX	17-SEP-2002	2002US-0410957P.
PR	17-SEP-2002	2002US-0410958P.
PR	17-SEP-2002	2002US-0410959P.
PR	17-SEP-2002	2002US-0410960P.
PR	17-SEP-2002	2002US-0410961P.
PR	17-SEP-2002	2002US-0410962P.
PR	17-SEP-2002	2002US-0410969P.
PR	17-SEP-2002	2002US-0411012P.
PR	17-SEP-2002	2002US-0411023P.
PR	17-SEP-2002	2002US-0411024P.
PR	17-SEP-2002	2002US-0411032P.
PR	17-SEP-2002	2002US-0411035P.
PR	17-SEP-2002	2002US-0411037P.
PR	17-SEP-2002	2002US-0411041P.
PR	17-SEP-2002	2002US-0411045P.
PR	17-SEP-2002	2002US-0411046P.
PR	17-SEP-2002	2002US-0411048P.
PR	17-SEP-2002	2002US-0411052P.
PR	17-SEP-2002	2002US-0411055P.
PR	17-SEP-2002	2002US-0411073P.
PR	17-SEP-2002	2002US-0411082P.
PR	17-SEP-2002	2002US-0411101P.
PR	17-SEP-2002	2002US-0411111P.
PR	18-APR-2003	2003US-0463700P.
PR	18-APR-2003	2003US-0463708P.
PR	18-APR-2003	2003US-0463716P.
PR	18-APR-2003	2003US-0463732P.
PR	02-MAY-2003	2003US-0467199P.
PR	02-MAY-2003	2003US-0467201P.
PR	02-MAY-2003	2003US-0467203P.
PR	02-MAY-2003	2003US-0467230P.
PR	19-MAY-2003	2003US-0471306P.
PR	19-MAY-2003	2003US-0471336P.
PR	22-MAY-2003	2003US-0472420P.
PR	22-MAY-2003	2003US-0472430P.
PR	09-JUN-2003	2003US-0476609P.
PR	09-JUN-2003	2003US-0476641P.
PR	08-JUL-2003	2003US-0485218P.
PR	08-JUL-2003	2003US-0485223P.
PR	08-JUL-2003	2003US-0485224P.
PR	08-JUL-2003	2003US-0485325P.
PR	14-JUL-2003	2003US-0486446P.
PR	14-JUL-2003	2003US-0486480P.
PR	15-JUL-2003	2003US-0486911P.
PR	15-JUL-2003	2003US-0486960P.
PR	08-AUG-2003	2003US-0493341P.
PR	08-AUG-2003	2003US-0493370P.
PR	08-AUG-2003	2003US-0493573P.
PR	08-AUG-2003	2003US-0493577P.
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
XX	Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D;	
PI	Hallenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;	
PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;	
XX	WPI, 2004-348438/32.	
DR		
XX	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	genetic, bacterial and viral diseases.	
XX		
XX	Claim 1; SEQ ID NO 2863; 428bp; English.	
XX		
CC	The present invention relates to an isolated nucleic acid molecule	
CC	encoding a polypeptide which is believed to be cytostatic,	
CC	anti-inflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	composition and methods are useful for diagnosing, preventing and	
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,	
CC	immune, metabolic, genetic, bacterial and viral diseases. The present	
CC	sequence represents a human secreted protein. The present sequence is	
CC	available on WIPOMB and is not in the specification.	
XX		

[illegible]

PR 17-SEP-2002; 2002US-041111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485233P.
PR 08-JUL-2003; 2003US-0485242P.
PR 08-JUL-2003; 2003US-0485255P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Williams LR, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisman L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.

PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.

PS Claim 1; SEQ ID NO 2938; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic.
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
XX available on WIPOMEB and is not in the specification.

XX Sequence 604 AA;

Query Match 69.2%; Score 18; DB 8; Length 604;

Best Local Similarity 20.0%; Pred. No. 4.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
DB 430 CTTATTAAAC 439

RESULT 45

ID ADP30941 standard; protein; 604 AA.

XX ADP30941;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #1708.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

XX WO2004035732-A2.

XX 29-APR-2004.

PF 28-AUG-2003; 2003WO-US026780.

PR 29-AUG-2002; 2002US-0405576P.

PR 29-AUG-2002; 2002US-0405579P.

PR 29-AUG-2002; 2002US-0405585P.

PR 29-AUG-2002; 2002US-0405588P.

PR 29-AUG-2002; 2002US-0405608P.

PR 29-AUG-2002; 2002US-0405611P.

PR 29-AUG-2002; 2002US-0405612P.

PR 29-AUG-2002; 2002US-0405616P.

PR 29-AUG-2002; 2002US-0405640P.

PR 29-AUG-2002; 2002US-0405642P.

PR 29-AUG-2002; 2002US-0405653P.

PR 29-AUG-2002; 2002US-0405655P.

PR 29-AUG-2002; 2002US-0405666P.

PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.

PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.

PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.

PR 17-SEP-2002; 2002US-0410961P.

PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.

PR 17-SEP-2002; 2002US-0411022P.

PR 17-SEP-2002; 2002US-0411023P.

PR 17-SEP-2002; 2002US-0411024P.

PR 17-SEP-2002; 2002US-0411032P.

PR 17-SEP-2002; 2002US-0411035P.

PR 17-SEP-2002; 2002US-0411037P.

PR 17-SEP-2002; 2002US-0411041P.

PR 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.

PR 17-SEP-2002; 2002US-0411048P.

PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.

PR 17-SEP-2002; 2002US-0411101P.

PR 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463700P.

PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463716P.

PR 18-APR-2003; 2003US-0463732P.

PR 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.

PR 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.

PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.

PR 09-JUN-2003; 2003US-0476641P.

PR 08-JUL-2003; 2003US-0485218P.

PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485242P.

PR 08-JUL-2003; 2003US-0485325P.

PR 14-JUL-2003; 2003US-0486446P.

PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kothakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2939; 428pp; English.
PS
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and antiviral. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 604 AA;

Query Match 69.2%; Score 18; DB 8; Length 604;
Best Local Similarity 20.0%; Pred. No. 4.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXC 10
Db 430 CTTATTAAAC 439

RESULT 46
ADP30507
ID ADP30507 standard; protein; 605 AA.
XX
AC ADP30507;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1274.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406662P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411013P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kothakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2505; 428pp; English.
PS
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,

CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.

XX
SQ Sequence 605 AA;

Query Match 69.2%; Score 18; DB 8; Length 605;
Best Local Similarity 20.0%; Pred. No. 4.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10
DB 570 CAATTATTC 579

RESULT 47

ID ADP31657 standard; protein; 617 AA.

XX ADP31657;

XX 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2424.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

OS WO2004035732-A2.

PN 29-APR-2004.

XX 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

XX 29-AUG-2002; 2002US-0406666P.

XX 17-SEP-2002; 2002US-0410946P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410957P.

XX 17-SEP-2002; 2002US-0410958P.

XX 17-SEP-2002; 2002US-0410959P.

XX 17-SEP-2002; 2002US-0410960P.

XX 17-SEP-2002; 2002US-0410961P.

XX 17-SEP-2002; 2002US-0410962P.

XX 17-SEP-2002; 2002US-0411019P.

XX 17-SEP-2002; 2002US-0411022P.

XX 17-SEP-2002; 2002US-0411023P.

XX 17-SEP-2002; 2002US-0411024P.

XX 17-SEP-2002; 2002US-0411032P.

XX 17-SEP-2002; 2002US-0411035P.

XX 17-SEP-2002; 2002US-0411037P.

XX 17-SEP-2002; 2002US-0411041P.

XX 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.

PR 17-SEP-2002; 2002US-0411048P.

PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.

PR 17-SEP-2002; 2002US-0411101P.

PR 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463700P.

PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463716P.

PR 18-APR-2003; 2003US-0463732P.

PR 02-MAY-2003; 2003US-0467119P.

PR 02-MAY-2003; 2003US-0467201P.

PR 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.

PR 09-JUN-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.

PR 08-JUL-2003; 2003US-0476641P.

PR 08-JUL-2003; 2003US-0485218P.

PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485224P.

PR 14-JUL-2003; 2003US-0485325P.

PR 14-JUL-2003; 2003US-0486464P.

PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486960P.

PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.

PR 08-AUG-2003; 2003US-0493577P.

XX (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;

PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

DR WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 3655; 428pp; English.

XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic.
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMEB and is not in the specification.

XX
SQ Sequence 617 AA;

Query Match 69.2%; Score 18; DB 8; Length 617;
Best Local Similarity 20.0%; Pred. No. 4.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10
DB 67 CAATTATTC 76

RESULT 48

ID ABR41596 standard; protein; 626 AA.

XX

AC ABR41596;
XX
XX 02-JUN-2003 (first entry)
XX
XX Human DITHP antigen recognition protein.
DE
XX
XX Human, dithp; diagnostic and therapeutic polynucleotide; diagnosis;
KM cancer; cell proliferative disorder; autoimmune disorder;
KM inflammatory disorder; infection; hormonal disorder; metabolic disorder;
KM neurological disorder; gastrointestinal disorder; transport disorder;
KM connective tissue disorder; drug screening; proteome analysis;
KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
KW disease model; toxicological testing; transcript imaging;
KW antigen recognition.
XX
XX Homo sapiens.
OS
XX WO200297031-A2.
XX
XX 05-DEC-2002.
XX
XX 27-MAR-2002; 2002WO-US010056.
XX
XX 28-MAR-2001; 2001US-0279619P.
PR 29-MAR-2001; 2001US-0280067P.
PR 29-MAR-2001; 2001US-0280068P.
PR 16-MAY-2001; 2001US-0291280P.
PR 17-MAY-2001; 2001US-0291829P.
PR 17-MAY-2001; 2001US-0291849P.
PR 19-JUN-2001; 2001US-0299428P.
PR 20-JUN-2001; 2001US-0299776P.
PR 20-JUN-2001; 2001US-0300001P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Thason O, Yap PE, Amshy SR;
PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kieselield Y, Gerstein EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Ian RV, Urashka ME;
XX
XX WPI: 2003-129518/12.
XX
XX N-PSDB; ACC46534.
DR
XX
XX Novel human diagnostic and therapeutic polypeptide useful for identifying
PT test compound which specifically binds to a polypeptide encoded by human
PI diagnostic and therapeutic polynucleotide, and to induce antibodies.
XX
PS Claim 27; SEQ ID NO 1131, 591pp; English.
XX
XX The invention relates to novel human diagnostic and therapeutic
CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to
CC polynucleotide sequences at least 90% identical to the dithp cDNA
CC sequences of the invention; recombinant vectors, host cells and
CC transgenic organisms comprising a dithp nucleic acid sequence; the
CC recombinant production of DITHP proteins; antibodies specific for DITHP
CC proteins; microarrays comprising dithp nucleic acid sequences; methods of
CC detecting dithp nucleotide and protein sequences; methods of screening
CC for compounds which specifically bind a DITHP protein; and methods of
CC assessing the toxicity of test compounds using a dithp hybridisation
CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
CC diagnosis of a wide variety of conditions including cancer and other cell
CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
CC viral, fungal or parasitic infections; hormonal disorders; metabolic
CC disorders; neurological disorders; gastrointestinal disorders; transport
CC disorders; and connective tissue disorders. They may also be used to
CC screen for modulators of protein activity or gene expression. DITHP
CC proteins can additionally be used in analysis of the proteome of a tissue
CC or cell type and to induce antibodies. The dithp nucleic acids are
CC additionally useful in somatic or germline gene therapy of the disorders
CC mentioned above, as a source of antisense sequences, as a source of
CC probes and primers, in genotyping and identification of individuals, in

CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
CC The present sequence represents a DITHP protein which has antigen
CC recognition activity. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 626 AA;
XX
XX Query Match 69.2%; Score 18; DB 6; Length 626;
XX Best Local Similarity 20.0%; Pred. No. 4.7;
XX Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
XX
QY 1 CXXXXXXXXXC 10
DB 41 CTSTSSAASC 50
XX
RESULT 49
ADP30858
ID ADP30858 standard; protein; 645 AA.
XX
XX ADP30858;
XX
XX 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #1625.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
OS
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.
 PR 17-SEP-2002; 2002US-0411048P.
 PR 17-SEP-2002; 2002US-0411052P.
 PR 17-SEP-2002; 2002US-0411055P.
 PR 17-SEP-2002; 2002US-0411073P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411101P.
 PR 18-APR-2003; 2003US-0463700P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 19-MAY-2003; 2003US-0467230P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 19-MAY-2003; 2003US-0471366P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 22-MAY-2003; 2003US-0472430P.
 PR 09-JUN-2003; 2003US-0476609P.
 PR 09-JUN-2003; 2003US-0476641P.
 PR 08-JUL-2003; 2003US-0485218P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485224P.
 PR 14-JUL-2003; 2003US-0486446P.
 PR 14-JUL-2003; 2003US-0486480P.
 PR 15-JUL-2003; 2003US-0486891P.
 PR 15-JUL-2003; 2003US-0486960P.
 PR 08-AUG-2003; 2003US-0493341P.
 PR 08-AUG-2003; 2003US-0493370P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

XX PA Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;
 PI Halebek RP, Huang MM, Kochakota S, Haislan L, Linnemann T;
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
 XX WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 2856; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule
 CC encoding a polypeptide which is believed to be cytostatic,
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is
 CC available on WIPOMB and is not in the specification.

XX Sequence 645 AA;

XX Query Match 69.2%; Score 18; DB 8; Length 645;

XX Best Local Similarity 20.0%; Pred. No. 4.7;
 XX Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10
 DB 161 CAAATTTTTC 170

RESULT 50
 AAAY94907
 ID AAAY94907 standard; protein; 653 AA.
 XX

AC AAAY94907;
 DT 16-JUN-2000 (first entry)

XX Human secreted protein clone ca106_19x protein sequence SEQ ID NO:20.

XX Human; secreted protein; immunostimulant; immunosuppressant; virucide;
 KW antibacterial; antifungal; cytostatic; antiinflammatory; dermatological;
 KW antidiabetic; antiaesthetic; antiarthritic; antirheumatic; protozoacide;
 KW antihydroid; immune deficiency; severe combined immunodeficiency; SCID;
 KW infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;
 KW connective tissue disease; multiple sclerosis; erythematosis;
 KW rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;
 KW Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;
 KW insulin dependent diabetes mellitus; graft-versus-host-disease;
 KW autoimmune inflammatory eye disease; allergy.

XX Homo sapiens.

OS WO200009552-A1.

XX 24-FEB-2000.

XX 13-AUG-1999; 99WO-US018298.

XX 14-AUG-1998; 98US-0096622P.

XX 17-AUG-1998; 98US-0096815P.

XX 04-SEP-1998; 98US-0099229P.

XX 23-OCT-1998; 98US-0105368P.

XX 08-JAN-1999; 99US-0115334P.

XX 12-FEB-1999; 99US-0119931P.

XX 18-FEB-1999; 99US-0120575P.

XX 30-APR-1999; 99US-0132020P.

XX 11-AUG-1999; 99US-0148424P.

XX (GENY) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
 PI Mong CG, Clark HF, Fechtel K;
 XX WPI; 2000-205979/18.

XX Claim 29; Page 487-489; 641pp; English.

XX AAA1618 to AAA1697 encode the human secreted proteins given in AAAY94898

CC to AAAY94980, isolated from human adult brain, adult thyroid, adult
 CC retina, foetal carcinoma, adult blood, adult neural, foetal kidney, adult
 CC placenta, adult testis, whole embryo, adult cartilage, kidney, foetal
 CC brain, adult thymus, foetal placenta, adult uterus, adult tumour, and
 CC adult bladder, cDNA libraries. The polynucleotides and proteins are
 CC predicted to have biological activities which would make them suitable
 CC for treating, preventing or ameliorating medical conditions in humans and
 CC animals. The polynucleotides can be used as markers for tissues in which
 CC the protein is preferentially expressed, as molecular weight markers on
 CC Southern gels, and as chromosome markers or tags to identify chromosomes
 CC or to map gene positions. The proteins can be used in the treatment of
 CC immune deficiencies and disorders, such as severe combined
 CC immunodeficiency (SCID), as well as viral, bacterial, fungal and other
 CC infections. These infections include human immunodeficiency virus (HIV),
 CC hepatitis, herpesviruses, mycobacteria, leishmania spp., malaria and
 CC candidiasis. The proteins can be used to treat autoimmune disorders such
 CC as connective tissue disease, multiple sclerosis, systemic lupus
 CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation,
 CC Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent
 CC diabetes mellitus, myasthenia gravis, graft-versus-host-disease and
 CC autoimmune inflammatory eye disease. The proteins can also be used to
 CC treat allergic conditions, such as asthma. AAA1698 to AAA16774 represent

CC probes for the human secreted proteins from the present invention
 XX
 SQ Sequence 653 AA;

Query Match 69.2%; Score 18; DB 3; Length 653;
 Best Local Similarity 20.0%; Pred. No. 4.7;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 10
 |
 Db 173 CAAAAAAC 182

Search completed: January 4, 2006, 15:56:17
 Job time : 70.4348 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 ; Search time 10.3478 Seconds
(without alignments)
92.983 Million cell updates/sec

Title: US-09-932-322-11
Perfect score: 26
Sequence: 1 CXXXXXXXXXC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 1000 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	69.2	313	2	S59448
2	18	69.2	480	2	E75433
3	17	65.4	194	2	JQ1560
4	17	65.4	220	2	JC4082
5	17	65.4	376	2	AB2920
6	17	65.4	388	2	C97694
7	17	65.4	389	2	T23167
8	17	65.4	570	2	T37314
9	17	65.4	634	2	S33575
10	17	65.4	942	2	D87803
11	17	65.4	2704	2	S09118
12	16	61.5	71	2	F84312
13	16	61.5	85	2	E70531
14	16	61.5	104	2	S51479
15	16	61.5	105	2	A71249
16	16	61.5	108	2	F72549
17	16	61.5	126	2	A23473
18	16	61.5	129	2	T49498
19	16	61.5	133	2	A96746
20	16	61.5	147	2	S09762
21	16	61.5	153	2	PN0103
22	16	61.5	161	2	T28088
23	16	61.5	164	2	JQ1252
24	16	61.5	229	2	T34277
25	16	61.5	242	2	A45724
26	16	61.5	244	2	S72219
27	16	61.5	251	2	A55035
28	16	61.5	262	1	JQ1724
29	16	61.5	263	2	S47537

30	16	61.5	275	2	AB5856	probable elongatio
31	16	61.5	275	2	B64986	hypothetical 30.9
32	16	61.5	275	2	G91011	probable elongatio
33	16	61.5	294	2	T23682	hypothetical prote
34	16	61.5	306	2	C71498	probable yop trans
35	16	61.5	317	2	T00500	probable alicitor
36	16	61.5	321	2	AB4792	hypothetical prote
37	16	61.5	349	2	JE0202	paired-box-contain
38	16	61.5	380	2	T04007	hypothetical prote
39	16	61.5	397	2	T26077	hypothetical prote
40	16	61.5	422	2	S41514	Rad52 protein homo
41	16	61.5	473	2	C81039	lipopolysaccharide
42	16	61.5	473	2	C81984	probable lipopolys
43	16	61.5	475	2	T36342	probable glutamate
44	16	61.5	486	2	B86460	hypothetical prote
45	16	61.5	510	2	A55207	glycerol-3-phospha
46	16	61.5	537	2	G31277	guinate transport
47	16	61.5	580	2	B38418	jockey protein 1 -
48	16	61.5	638	2	T22518	hypothetical prote
49	16	61.5	642	2	C81345	hypothetical prote
50	16	61.5	648	2	T37581	probable serine-ch
51	16	61.5	708	1	TEPOL	lactotransferrin p
52	16	61.5	708	2	JC2323	lactoferrin - goat
53	16	61.5	711	1	TEFUL	lactotransferrin p
54	16	61.5	907	2	AB6460	99.9K hypothetical
55	16	61.5	975	2	T29908	hypothetical prote
56	16	61.5	1031	2	T06130	hypothetical prote
57	16	61.5	1069	2	T22138	hypothetical prote
58	16	61.5	1105	2	T22132	hypothetical prote
59	16	61.5	1136	2	AB1581	different protein
60	16	61.5	1380	2	T18309	receptor-adenylate
61	16	61.5	1474	2	D88550	protein ZC84.6 [lm
62	16	61.5	2233	2	T28669	surface protein 51
63	16	61.5	2395	1	S50820	surface protein ty
64	16	61.5	2533	2	T28675	alpha-51D-immobil
65	16	61.5	2533	2	T28674	surface antigen - p
66	16	61.5	2543	2	T31687	G surface protein
67	16	61.5	2718	2	A23475	metallothionein Mc
68	15	57.7	40	1	SMFP	metallothionein Mc
69	15	57.7	40	2	B61194	hypothetical prote
70	15	57.7	43	2	S34483	hypothetical prote
71	15	57.7	74	2	AB2210	hypothetical prote
72	15	57.7	93	2	T18118	hypothetical prote
73	15	57.7	109	2	S64309	hypothetical prote
74	15	57.7	111	2	A72079	hypothetical prote
75	15	57.7	111	2	B86544	hypothetical prote
76	15	57.7	113	2	T10136	hypothetical prote
77	15	57.7	123	2	B81120	probable periplasm
78	15	57.7	124	2	C81707	hypothetical prote
79	15	57.7	131	2	T36866	probable iron-sulf
80	15	57.7	135	2	T38956	very hypothetical
81	15	57.7	139	2	D97348	hypothetical prote
82	15	57.7	141	2	AE0029	hypothetical prote
83	15	57.7	141	2	PCI294	MeR-family trans
84	15	57.7	146	2	T01683	crophosolite surfac
85	15	57.7	149	2	AR2977	immobilization ant
86	15	57.7	158	2	T14975	conserved hypothet
87	15	57.7	160	2	F98305	hypothetical prote
88	15	57.7	169	1	S18946	hypothetical 16.6K
89	15	57.7	171	2	T10315	ultra high-sulfur
90	15	57.7	179	2	T22471	hypothetical prote
91	15	57.7	181	2	T31481	hypothetical prote
92	15	57.7	185	2	T51003	hypothetical prote
93	15	57.7	189	2	B90969	probable tail asse
94	15	57.7	191	2	S69735	hypothetical prote
95	15	57.7	195	2	T18617	hypothetical prote
96	15	57.7	199	2	C90734	probable outer mem
97	15	57.7	199	2	B85584	hypothetical prote
98	15	57.7	203	2	H81695	recombination prot
99	15	57.7	204	2	A28864	ShuY-like protein
100	15	57.7	207	2	A91177	hypothetical prote
101	15	57.7	207	2	B86023	hypothetical-relat
102	15	57.7	208	2	T07732	tuberculosis-relat

103	15	57.7	211	2	S34274	probable secretory
104	15	57.7	211	2	UC4883	scytalidopepsin B
105	15	57.7	212	2	T22463	hypothetical prote
106	15	57.7	214	2	C70812	probable 1pqq prot
107	15	57.7	219	1	QRECTC	carbonate dehydrat
108	15	57.7	219	2	H90677	carbonic anhydrase
109	15	57.7	219	2	D85528	carbonic anhydrase
110	15	57.7	220	2	S22493	coat protein - Cym
111	15	57.7	223	2	T33194	hypothetical prote
112	15	57.7	223	2	T33193	hypothetical prote
113	15	57.7	227	2	C90899	probable tail asse
114	15	57.7	230	2	T49555	related to high cy
115	15	57.7	231	2	S46511	isopenyl transfer
116	15	57.7	231	2	S41894	gene 3' protein -
117	15	57.7	238	1	JH0367	ribonuclease all160
118	15	57.7	242	1	AB2007	transposase all160
119	15	57.7	244	2	T31838	hypothetical prote
120	15	57.7	245	1	KY80A	chymotrypsin (EC 3
121	15	57.7	245	1	KY80B	chymotrypsin (EC 3
122	15	57.7	247	2	B90997	probable tail asse
123	15	57.7	247	2	T27778	hypothetical prote
124	15	57.7	248	2	T03868	hypothetical prote
125	15	57.7	248	2	T03869	hypothetical prote
126	15	57.7	248	2	T31841	hypothetical prote
127	15	57.7	256	1	TRRF	trypsin-like prote
128	15	57.7	256	1	T27836	hypothetical prote
129	15	57.7	258	2	C86286	protein P9L1.16 [1
130	15	57.7	260	2	E90344	hydroxyethylthiazol
131	15	57.7	262	1	G64977	hydroxyethylthiazol
132	15	57.7	262	2	E85837	hydroxyethylthiazol
133	15	57.7	262	2	C90992	hydooxyethylthiazol
134	15	57.7	263	1	KYR1B	chymotrypsin (EC 3
135	15	57.7	263	2	A31299	chymotrypsin (EC 3
136	15	57.7	263	2	A21195	chymotrypsin (EC 3
137	15	57.7	264	2	T38136	unknown protein T3
138	15	57.7	264	2	F96807	hypothetical prote
139	15	57.7	273	2	A05113	translacton elonga
140	15	57.7	274	2	S64523	hypothetical prote
141	15	57.7	293	2	A85817	probable tail comp
142	15	57.7	293	2	E85718	hypothetical prote
143	15	57.7	293	2	T31840	hypothetical prote
144	15	57.7	294	2	E84640	CONSTANS-like B-po
145	15	57.7	297	2	H72779	hypothetical prote
146	15	57.7	307	2	T08125	CONSTANS protein h
147	15	57.7	307	2	T45213	superoxide dismuta
148	15	57.7	311	2	T15268	hypothetical prote
149	15	57.7	312	2	JCS522	uncoupling protein
150	15	57.7	312	2	T31834	hypothetical prote
151	15	57.7	323	2	S72915	hypothetical prote
152	15	57.7	326	2	T19996	hypothetical prote
153	15	57.7	337	2	T23794	hypothetical prote
154	15	57.7	344	2	S42383	hypothetical prote
155	15	57.7	345	2	A05279	surface antigen 51
156	15	57.7	363	2	JCS536	C-Fringe protein 1
157	15	57.7	368	2	S18758	hypothetical prote
158	15	57.7	379	2	B84073	sulfate adenyllylr
159	15	57.7	382	2	T27058	hypothetical prote
160	15	57.7	385	2	T25492	hypothetical prote
161	15	57.7	386	2	T32240	hypothetical prote
162	15	57.7	390	2	C95954	hypothetical prote
163	15	57.7	392	2	S65693	opioid receptor mu
164	15	57.7	395	1	A35629	mevalonate kinase
165	15	57.7	400	2	I56553	mu opiate receptor
166	15	57.7	410	2	T39115	formamide-like p
167	15	57.7	412	2	AF0668	probable benzate
168	15	57.7	423	1	TVPF2E	transforming prote
169	15	57.7	423	2	S38953	carboxypeptidase D
170	15	57.7	423	2	A41204	carboxypeptidase B
171	15	57.7	450	1	S13730	pmba protein - Bsc
172	15	57.7	450	2	AB1056	probable PmbA prot
173	15	57.7	450	2	D86121	maturation of anti
174	15	57.7	450	2	D91280	maturation of anti
175	15	57.7	454	2	T34297	hypothetical prote
176	15	57.7	459	2	T48854	gene murine tumour
177	15	57.7	461	1	A35356	tumor necrosis fac
178	15	57.7	470	2	A40697	63k sperm flagella
179	15	57.7	471	2	T39571	probable regulator
180	15	57.7	474	2	B38634	tumor necrosis fac
181	15	57.7	474	2	T43504	sp06 protein - fis
182	15	57.7	476	2	T29054	probable transmemb
183	15	57.7	482	2	S65766	G protein-coupled
184	15	57.7	488	2	T10794	pherophorin III -
185	15	57.7	493	2	S46826	hypothetical prote
186	15	57.7	493	2	S36931	chitinase (EC 3.2.
187	15	57.7	495	2	S31493	env polypeptid -
188	15	57.7	503	2	S59698	HST1 protein - yea
189	15	57.7	504	2	C85485	probable carnitine
190	15	57.7	504	2	C90634	probable carnitine
191	15	57.7	504	2	H64724	probable carnitine
192	15	57.7	505	2	AD0511	probable carnitine
193	15	57.7	505	2	T19971	hypothetical prote
194	15	57.7	506	2	T19973	hypothetical prote
195	15	57.7	512	2	S28267	pherophorin I prec
196	15	57.7	512	2	T06585	ammonium transport
197	15	57.7	525	2	F84933	IMP cyclohydrolase
198	15	57.7	526	2	T07082	lycopene epsilon-c
199	15	57.7	529	2	S51477	extracellular glyc
200	15	57.7	529	2	T48253	myb-like protein -
201	15	57.7	533	2	T00742	ubiquitin-binding
202	15	57.7	543	2	S65462	glucose transport
203	15	57.7	545	2	A46281	tetrahydrofolylpol
204	15	57.7	546	2	T11217	reverse transcript
205	15	57.7	549	2	JN0553	triacylglycerol li
206	15	57.7	551	2	B88949	protein R0985.4 [1
207	15	57.7	553	2	T01479	hypothetical prote
208	15	57.7	561	2	T27318	hypothetical prote
209	15	57.7	562	1	RGBYS2	regulatory protein
210	15	57.7	568	2	T31692	hypothetical prote
211	15	57.7	571	2	T52325	pectinesterase (EC
212	15	57.7	576	2	H87414	TPR domain protein
213	15	57.7	584	2	S14952	pectinesterase hom
214	15	57.7	586	2	T52330	pectinesterase (EC
215	15	57.7	588	2	T02184	probable pectinest
216	15	57.7	595	2	T52327	pectinesterase (EC
217	15	57.7	595	2	T02183	probable pectinest
218	15	57.7	599	2	T10798	pherophorin-S - Vo
219	15	57.7	616	2	T29234	hypothetical prote
220	15	57.7	626	2	T27319	hypothetical prote
221	15	57.7	634	1	H64390	carbon-monooxide de
222	15	57.7	641	2	T27959	hypothetical prote
223	15	57.7	642	2	T05497	hypothetical prote
224	15	57.7	646	2	T35002	probable respirato
225	15	57.7	661	2	E71427	hypothetical prote
226	15	57.7	670	2	S36616	regulatory protein
227	15	57.7	676	2	T30480	envelope protein h
228	15	57.7	698	1	TFHUP	transferrin precu
229	15	57.7	698	2	D90771	hypothetical prote
230	15	57.7	698	2	H85633	hypothetical prote
231	15	57.7	698	2	B54839	ye5H protein precu
232	15	57.7	700	2	T32205	hypothetical prote
233	15	57.7	703	2	A45543	lactoferrin precu
234	15	57.7	713	2	A35502	major surface-labe
235	15	57.7	716	2	T09462	juvenile hormone r
236	15	57.7	717	2	S12100	transferrin precu
237	15	57.7	717	2	T25431	hypothetical prote
238	15	57.7	733	2	A35421	probable oxidoredu
239	15	57.7	738	2	T26421	hypothetical prote
240	15	57.7	739	2	B88553	hypothetical prote
241	15	57.7	744	2	B85092	protein K04H4.2b [
242	15	57.7	748	2	S57061	nuclear pore prote
243	15	57.7	748	2	B45046	basic juvenile hor
244	15	57.7	749	2	B35889	probable dehydroge
245	15	57.7	750	2	H91074	transcription regu
246	15	57.7	750	2	G85919	transcription regu
247	15	57.7	750	2	D65051	hyPF protein - Bsc
248	15	57.7	751	2	F87789	protein C34G6.2 [1

249	15	57.7	752	2	G85941	hypothetical prote	322	15	57.7	2090	2	T30075	hypothetical prote
250	15	57.7	752	2	C91096	hypothetical prote	323	15	57.7	2153	2	T30074	hypothetical prote
251	15	57.7	752	2	B65070	probable oxidoredu	324	15	57.7	2201	2	A54774	ATP binding casset
252	15	57.7	755	2	A44315	cartilage oligomer	325	15	57.7	2476	2	T34022	zonadhesin - pig
253	15	57.7	771	1	WZB52	gene 52 protein -	326	15	57.7	2531	2	T31070	notch homolog - se
254	15	57.7	793	1	KXMSF	furin (BC 3.4.21.7	327	15	57.7	2844	2	S28291	hypothetical prote
255	15	57.7	793	1	KXRTF	furin (BC 3.4.21.7	328	15	57.7	2910	2	T42214	otogellin - mouse
256	15	57.7	794	1	KXRTF	furin (BC 3.4.21.7	329	15	57.7	2918	2	A54105	fibillin-2 precur
257	15	57.7	797	1	I46044	furin (BC 3.4.21.7	330	15	57.7	3097	2	T00021	DN-cadherin - fru
258	15	57.7	806	2	T18840	hypothetical prote	331	15	57.7	3147	2	T21328	hypothetical prote
259	15	57.7	814	2	T49207	receptor kinase-11	332	15	57.7	3712	2	S18253	laminin alpha-1 ch
260	15	57.7	846	1	VCLJND	env polypeptide pr	333	15	57.7	4660	2	T42737	gp330 protein prec
261	15	57.7	847	1	S36337	histidine decarbox	334	15	57.7	4957	2	T03455	ALR protein - huma
262	15	57.7	849	2	T49342	fzr related protei	335	15	57.7	5262	2	T03454	subtilisin A [vali
263	15	57.7	852	2	S41886	DNA repair protein	336	15	53.8	43	2	A69704	fulvicin C - Myxoc
264	15	57.7	854	2	S54384	envelope polypept	337	14	53.8	45	1	FRYZ	calitoxin - sea an
265	15	57.7	854	1	VCLJST	env polypeptide pr	338	14	53.8	46	2	A31863	viral protein - Ag
266	15	57.7	855	1	VCLJZR	env polypeptide pr	339	14	53.8	55	1	B7AGA6	lipoprotein R21 pr
267	15	57.7	856	1	A44963	gag polypeptide -	340	14	53.8	60	2	JN0750	lipoprotein R21 pr
268	15	57.7	863	2	A53034	hypothetical prote	341	14	53.8	60	2	H90831	metallothionein-2a
269	15	57.7	863	2	T27958	hypothetical prote	342	14	53.8	62	2	S54336	metallothionein-2c
270	15	57.7	873	2	B87049	conserved hypothet	343	14	53.8	62	2	S54335	hypothetical prote
271	15	57.7	888	1	GMLJND	pol polypeptide -	344	14	53.8	62	2	T10302	hypothetical prote
272	15	57.7	890	2	T21000	hypothetical prote	345	14	53.8	65	2	S77379	hypothetical prote
273	15	57.7	891	2	E96590	hypothetical prote	346	14	53.8	73	2	F70641	hypothetical prote
274	15	57.7	895	2	B96775	hypothetical prote	347	14	53.8	73	2	G97142	hypothetical prote
275	15	57.7	911	2	S46497	aspartate kinase (348	14	53.8	75	2	A10863	probable lipoprote
276	15	57.7	911	2	E86438	hypothetical prote	349	14	53.8	76	2	E91087	hypothetical prote
277	15	57.7	914	2	T17233	hypothetical prote	350	14	53.8	76	2	E65063	hypothetical prote
278	15	57.7	915	2	T21773	hypothetical prote	351	14	53.8	77	2	T10276	hypothetical prote
279	15	57.7	928	2	T20035	glycoprotein GP330	352	14	53.8	78	1	IKEC51	colicin V immunity
280	15	57.7	928	2	A30363	hypothetical prote	353	14	53.8	80	2	PM0044	estrogen receptor
281	15	57.7	975	2	T48107	hypothetical prote	354	14	53.8	83	2	H81188	hypothetical prote
282	15	57.7	978	2	H86319	hypothetical prote	355	14	53.8	85	2	E83008	hypothetical prote
283	15	57.7	982	1	VCLJLK	env polypeptide -	356	14	53.8	86	2	S78488	hypothetical prote
284	15	57.7	995	2	S50358	hypothetical prote	357	14	53.8	86	2	T07829	Ig kappa chain V r
285	15	57.7	1012	1	DJB865	DNA-directed DNA p	358	14	53.8	88	2	JC1126	systemic acquired
286	15	57.7	1012	2	T44185	probable DNA-direc	359	14	53.8	91	2	J50036	major allergen cha
287	15	57.7	1012	2	T43998	DNA polymerase [lm	360	14	53.8	91	2	B87324	Clara cell 10K pro
288	15	57.7	1012	2	T41940	DNA polymerase - h	361	14	53.8	91	2	T37327	hypothetical prote
289	15	57.7	1013	2	QJ1920	DNA-directed DNA p	362	14	53.8	92	1	S36358	insulin homolog ce
290	15	57.7	1014	2	T30431	DNA-directed DNA p	363	14	53.8	92	1	UGMS	proteinase inhibit
291	15	57.7	1018	2	T19693	hypothetical prote	364	14	53.8	96	2	C71614	ribosomal protein
292	15	57.7	1047	2	S19508	hypothetical prote	365	14	53.8	96	2	A36581	polyoxorinated bi
293	15	57.7	1059	2	T22545	MSH3 protein - yea	366	14	53.8	97	2	S72866	hypothetical prote
294	15	57.7	1063	2	T46284	hypothetical prote	367	14	53.8	98	2	A10334	conserved hypothet
295	15	57.7	1064	2	A40136	fibropellin 1a - s	368	14	53.8	99	2	JC2136	monocyte chemoattr
296	15	57.7	1083	2	S59780	hypothetical prote	369	14	53.8	99	2	S60230	glibetellin-regula
297	15	57.7	1095	2	T13964	probable histone d	370	14	53.8	100	2	S64316	hypothetical prote
298	15	57.7	1096	2	T16875	hypothetical prote	371	14	53.8	102	2	F70534	hypothetical prote
299	15	57.7	1118	2	A48292	mucin, tracheobron	372	14	53.8	103	2	H84599	hypothetical prote
300	15	57.7	1136	2	AH1227	different proteins	373	14	53.8	103	2	S70182	hypothetical prote
301	15	57.7	1193	2	T50729	magnesium-protopor	374	14	53.8	105	2	S42214	hydroxymethylgluta
302	15	57.7	1224	2	T07446	DNA-directed RNA p	375	14	53.8	105	2	S42212	hydroxymethylgluta
303	15	57.7	1245	2	D86260	protein T12C24.22	376	14	53.8	105	2	S17345	hydroxymethylgluta
304	15	57.7	1274	2	T10729	transferrin-like p	377	14	53.8	105	2	S42213	hydroxymethylgluta
305	15	57.7	1274	2	T42017	cysteine rich prot	378	14	53.8	105	2	S42215	pectin esterase ho
306	15	57.7	1284	2	S06187	RNA2 polypeptide -	379	14	53.8	105	2	T14406	galactose-binding
307	15	57.7	1357	2	B96696	protein FN21.4 (1	380	14	53.8	105	2	A37961	zinc finger protei
308	15	57.7	1391	2	T20406	hypothetical prote	381	14	53.8	105	2	B48827	hypothetical prote
309	15	57.7	1477	2	T18534	protein-tyrosine k	382	14	53.8	106	4	S57386	hypothetical prote
310	15	57.7	1524	2	T30337	polypeptide - Afri	383	14	53.8	107	1	T49527	putidaredoxin [val
311	15	57.7	1544	2	T04464	hypothetical prote	384	14	53.8	107	2	T49527	hypothetical prote
312	15	57.7	1584	2	T00026	brain-specific ang	385	14	53.8	109	2	E72850	Acotf-5 protein
313	15	57.7	1626	2	T26318	hypothetical prote	386	14	53.8	109	2	T41895	orfs homolog orf13
314	15	57.7	1687	2	T30176	BGF repeat: transme	387	14	53.8	109	2	A13178	conserved hypothet
315	15	57.7	1722	2	E89753	protein F11C7.4 (1	388	14	53.8	111	2	S64475	hypothetical prote
316	15	57.7	1743	2	T26859	hypothetical prote	389	14	53.8	112	2	S642929	hypothetical prote
317	15	57.7	1820	2	A55494	latent transformat	390	14	53.8	113	2	S07216	gonadotropin I bet
318	15	57.7	1827	2	A35694	cut1 protein - fis	391	14	53.8	113	2	T20299	hypothetical prote
319	15	57.7	1828	2	T41455	cut1 protein - fis	392	14	53.8	116	1	S15197	hypA protein - Bsc
320	15	57.7	1863	1	A58881	breast/ovarian can	393	14	53.8	116	2	F85921	pleiotropic effec
321	15	57.7	1866	1	GNWB2C	genome polypeptide	394	14	53.8	116	2	F91076	HypA protein (limpo

395	14	53.8	116	2	A54598	unilateral minicirc
396	14	53.8	116	2	D95280	hypothetical prote
397	14	53.8	116	2	F81285	probable periplasm
398	14	53.8	118	1	PSSNK1	phospholipase A2 (
399	14	53.8	118	1	PSSNK3	phospholipase A2 (
400	14	53.8	118	2	D34860	phospholipase A2 (
401	14	53.8	118	2	E34860	phospholipase A2 (
402	14	53.8	118	2	G34860	phospholipase A2 (
403	14	53.8	118	2	F34860	phospholipase A2 (
404	14	53.8	118	2	H34860	phospholipase A2 (
405	14	53.8	118	2	AB0847	Hypa protein (impo
406	14	53.8	119	2	AC3199	transcription regu
407	14	53.8	120	2	F70424	pilin - Aquifex ae
408	14	53.8	121	2	H71351	probable ribosomal
409	14	53.8	123	2	T04635	hypothetical prote
410	14	53.8	124	2	E70754	hypothetical prote
411	14	53.8	125	2	AF2154	hypothetical prote
412	14	53.8	127	2	F70910	hypothetical prote
413	14	53.8	128	2	D69326	conserved hypothet
414	14	53.8	129	2	AH3212	hypothetical prote
415	14	53.8	131	1	BGB02	spermatid transiti
416	14	53.8	133	2	B61108	gonadotropin I bet
417	14	53.8	137	2	B36179	gonadotropin I bet
418	14	53.8	137	2	I51231	gonadotropin I bet
419	14	53.8	137	2	S34349	gonadotropin I bet
420	14	53.8	137	2	S55364	serine proteinase
421	14	53.8	138	2	T21792	hypothetical prote
422	14	53.8	139	2	S69458	hypothetical prote
423	14	53.8	139	2	PC4217	hypothetical 139 p
424	14	53.8	139	2	T12618	homobox protein H
425	14	53.8	140	2	T04904	hypothetical prote
426	14	53.8	141	2	A85994	probable transcrip
427	14	53.8	141	2	AE1009	probable Zn(II)-re
428	14	53.8	141	2	E91148	zinc (II) responsi
429	14	53.8	141	2	I67892	embryonic abundant
430	14	53.8	141	2	T09251	probable CO-induce
431	14	53.8	142	2	T51316	ORF MSV132 probabl
432	14	53.8	142	2	T28293	hypothetical prote
433	14	53.8	142	2	E84040	hypothetical prote
434	14	53.8	142	2	JS0510	fusaric acid resis
435	14	53.8	142	2	T03976	hypothetical prote
436	14	53.8	142	4	S13768	MHC class I histoc
437	14	53.8	143	2	H83935	hypothetical prote
438	14	53.8	143	2	T00293	hypothetical prote
439	14	53.8	143	2	T27692	hypothetical prote
440	14	53.8	144	2	S01391	phospholipase A2 (
441	14	53.8	145	2	T05641	hypothetical prote
442	14	53.8	146	2	T25356	hypothetical prote
443	14	53.8	146	2	C86187	YUPBH12.12 (impor
444	14	53.8	146	2	AD3427	transposase BME114
445	14	53.8	147	2	S42552	proline-rich prote
446	14	53.8	148	2	T30627	hypothetical prote
447	14	53.8	148	2	T48981	hypothetical prote
448	14	53.8	148	2	B72513	hypothetical prote
449	14	53.8	150	2	S11441	plasmin C precursor
450	14	53.8	151	2	B83550	positive regulator
451	14	53.8	151	2	T09256	heat shock protein
452	14	53.8	152	1	KRSHHC	keratin high-sulfu
453	14	53.8	152	2	I47111	high-sulfur wool m
454	14	53.8	152	2	I47109	high-sulfur wool m
455	14	53.8	152	2	I47108	high-sulfur wool m
456	14	53.8	152	2	I47112	high-sulfur wool m
457	14	53.8	152	2	A84503	hypothetical prote
458	14	53.8	152	2	T09253	heat shock protein
459	14	53.8	152	2	D89753	protein P11C7.2 (l
460	14	53.8	153	2	F75514	conserved hypothet
461	14	53.8	153	2	T32950	hypothetical prote
462	14	53.8	154	1	QQVLAW	gene X protein - h
463	14	53.8	154	1	QQVLCP	gene X protein - h
464	14	53.8	154	1	QQVLKS	gene X protein - h
465	14	53.8	154	2	S47408	gene X protein - h
466	14	53.8	154	2	S20756	gene X protein - h
467	14	53.8	154	2	S67503	gene X protein - h
468	14	53.8	154	2	S12542	gene X protein - h
469	14	53.8	154	2	S33687	gene X protein - h
470	14	53.8	155	2	D72761	hypothetical prote
471	14	53.8	156	2	I67751	dopamine receptor
472	14	53.8	156	2	B83185	probable transcrip
473	14	53.8	156	2	B82073	sigma-E factor reg
474	14	53.8	159	1	I83299	sigma-E factor reg
475	14	53.8	159	2	AF0829	sigma-E factor reg
476	14	53.8	159	2	B85903	sigma-E factor, ne
477	14	53.8	159	2	D91058	sigma-E factor reg
478	14	53.8	159	2	T02999	sigma-E factor reg
479	14	53.8	160	2	T28953	hypothetical prote
480	14	53.8	160	2	JC7606	globin chain al pr
481	14	53.8	161	2	T21405	hypothetical prote
482	14	53.8	162	2	A87755	protein T21B2.3 (
483	14	53.8	166	2	E71342	probable deoxycyti
484	14	53.8	167	2	D81049	hypothetical prote
485	14	53.8	171	2	S09903	hypothetical prote
486	14	53.8	171	2	S09759	hypothetical prote
487	14	53.8	171	2	H83360	probable acetyltra
488	14	53.8	172	2	B81828	hypothetical integ
489	14	53.8	173	2	A83450	invasion protein b
490	14	53.8	173	2	A81923	probable membrane
491	14	53.8	174	1	JQ1625	small hydrophobic
492	14	53.8	176	2	T28762	hypothetical prote
493	14	53.8	177	2	B81085	conserved hypotet
494	14	53.8	177	2	S33166	protective antigen
495	14	53.8	177	2	H95981	conserved hypotet
496	14	53.8	178	2	A81687	conserved hypotet
497	14	53.8	178	2	S64257	hypothetical prote
498	14	53.8	181	2	T78572	Est3 protein - yea
499	14	53.8	182	2	T17625	hypothetical prote
500	14	53.8	184	2	D83234	hypothetical prote
501	14	53.8	185	2	T23777	hypothetical prote
502	14	53.8	187	2	T09390	21K protein precu
503	14	53.8	189	2	T28554	hypothetical prote
504	14	53.8	189	2	D36849	Al2u protein - var
505	14	53.8	189	2	B72165	Al3u protein - var
506	14	53.8	189	2	G95514	hypothetical prote
507	14	53.8	189	2	JK0235	core protein MGC-2
508	14	53.8	191	2	QJ1999	hypothetical 21.5K
509	14	53.8	191	2	T32278	hypothetical prote
510	14	53.8	192	2	E42518	Al2u protein - vac
511	14	53.8	193	2	D83296	hypothetical prote
512	14	53.8	196	2	T25943	hypothetical prote
513	14	53.8	197	2	B85068	hypothetical prote
514	14	53.8	197	2	T26829	hypothetical prote
515	14	53.8	197	2	C88809	protein Y43C5B.1 (
516	14	53.8	197	2	S72709	Leb1170 C3 229 pr
517	14	53.8	198	2	S72709	ORF1 protein - equ
518	14	53.8	198	2	A16646	hypothetical prote
519	14	53.8	200	2	B83642	resolvase - Klebsi
520	14	53.8	201	2	S38652	hypothetical prote
521	14	53.8	201	2	D82779	hypothetical prote
522	14	53.8	202	2	E70150	hypothetical prote
523	14	53.8	203	2	T02386	hypothetical prote
524	14	53.8	205	2	T02385	hypothetical prote
525	14	53.8	205	2	T04918	hypothetical prote
526	14	53.8	207	2	T22436	hypothetical prote
527	14	53.8	207	2	S70533	hypothetical prote
528	14	53.8	208	2	B75458	hypothetical prote
529	14	53.8	208	2	P00278	hypothetical prote
530	14	53.8	209	2	C70857	probable Igpa prot
531	14	53.8	211	2	S25938	env protein - huma
532	14	53.8	211	2	T18799	hypothetical prote
533	14	53.8	212	2	S09623	agglutinin isolect
534	14	53.8	213	1	AEW72	hypothetical prote
535	14	53.8	213	2	T27841	conserved hypotet
536	14	53.8	213	2	B82698	hemiferrin - bovin
537	14	53.8	215	2	A60166	hemiferrin - bovin
538	14	53.8	215	2	A71541	hemiferrin - bovin
539	14	53.8	216	1	A39684	hemiferrin - rat
540	14	53.8	216	2	H88102	protein W10G11.13

541	14	53.8	216	2	S60048	chlorophyll a/c-bi	614	14	53.8	273	2	G83403	conserved hypochet
542	14	53.8	217	2	T25618	hypothetical prote	615	14	53.8	274	2	UC4171	trypcase (EC 3.4.2
543	14	53.8	217	2	D69428	hypothetical prote	616	14	53.8	274	2	AB3112	transcription regu
544	14	53.8	218	2	T34373	hypothetical prote	617	14	53.8	274	2	H98174	transcription regu
545	14	53.8	219	2	S25939	env protein - huma	618	14	53.8	275	2	A32410	trypcase (EC 3.4.2
546	14	53.8	220	2	D83390	carbonate dehydrat	619	14	53.8	275	2	G83522	conserved hypochet
547	14	53.8	222	1	MMVZB4	antithrombin-III h	620	14	53.8	275	2	T29952	hypothetical prote
548	14	53.8	223	2	A65172	protein T08A9.3 [i	621	14	53.8	277	2	S71222	xyloglucan endo-1,
549	14	53.8	223	2	P89567	probable peroxisom	622	14	53.8	280	2	D97185	xyloglucan endo-1,
550	14	53.8	223	2	T37974	hypothetical prote	623	14	53.8	280	2	G01884	dUDP-4-dehydroxam
551	14	53.8	224	2	T01325	hypothetical prote	624	14	53.8	280	2	P95998	LIM protein FHL-1,
552	14	53.8	225	2	C86563	CT465 hypothetical	625	14	53.8	281	2	AE2248	probable sugar nuc
553	14	53.8	225	2	C72062	conserved hypochet	626	14	53.8	281	2	C86221	phosphomethylprim
554	14	53.8	226	2	T27843	hypothetical prote	627	14	53.8	282	2	T15556	hypothetical prote
555	14	53.8	228	2	T20219	hypothetical prote	628	14	53.8	282	2	T15640	hypothetical prote
556	14	53.8	229	2	T27840	hypothetical prote	629	14	53.8	283	2	T18969	5'-methylthiodeno
557	14	53.8	229	2	B82544	hypothetical prote	630	14	53.8	283	2	T21622	hypothetical prote
558	14	53.8	231	2	S28703	hypothetical prote	631	14	53.8	284	2	H97220	transcriptional reg
559	14	53.8	231	2	F64642	hypothetical prote	632	14	53.8	284	2	T06200	xyloglucan endo-1,
560	14	53.8	231	2	T24669	hypothetical prote	633	14	53.8	284	2	S04278	hypoxanthine phosp
561	14	53.8	232	2	E71325	probable V-type AT	634	14	53.8	285	1	IDBCRP	replication initia
562	14	53.8	234	2	T32843	hypothetical prote	635	14	53.8	285	2	T64780	replication-associ
563	14	53.8	237	2	T28070	hypothetical prote	636	14	53.8	285	2	T00306	replication initia
564	14	53.8	238	2	T04166	thiamatin-like pro	637	14	53.8	285	2	B88546	protein R107.2 [im
565	14	53.8	238	2	T40568	hypothetical prote	638	14	53.8	285	2	T24665	hypoxanthine phosp
566	14	53.8	239	2	B81345	probable periplasm	639	14	53.8	285	2	S09614	hypoxanthine phosp
567	14	53.8	240	2	T20319	hypothetical prote	640	14	53.8	286	2	S30872	hypothetical prote
568	14	53.8	242	2	T29854	hypothetical prote	641	14	53.8	286	2	B36570	ethanolamine ammon
569	14	53.8	243	2	A33329	hypothetical prote	642	14	53.8	286	2	T23354	hypothetical prote
570	14	53.8	243	2	JB0204	testis-specific pr	643	14	53.8	286	2	T31494	hypothetical prote
571	14	53.8	244	1	UC2018	phosducin - mouse	644	14	53.8	287	2	A85059	hypothetical prote
572	14	53.8	244	1	T37493	phosducin - cat	645	14	53.8	288	2	T02546	hypothetical prote
573	14	53.8	245	2	S52096	phosducin - cat	646	14	53.8	288	2	T21790	hypothetical prote
574	14	53.8	245	2	AP0412	sugar fermentation	647	14	53.8	289	2	T34688	probable lipoprote
575	14	53.8	245	2	T24565	hypothetical prote	648	14	53.8	290	2	C71321	hypothetical prote
576	14	53.8	246	1	A35422	phosducin, retinal	649	14	53.8	290	2	T10104	matrase-like prote
577	14	53.8	247	1	T43324	synaptogyrin homol	650	14	53.8	292	2	T19669	hypothetical prote
578	14	53.8	248	2	S49323	chymotrypsin (EC 3	651	14	53.8	292	2	G88783	protein C33A12.16
579	14	53.8	248	2	T21786	hypothetical prote	652	14	53.8	293	2	B49539	xyloglucan endo-1,
580	14	53.8	248	2	T22769	hypothetical prote	653	14	53.8	294	2	A48844	TGF alpha-like pro
581	14	53.8	249	2	T31837	hypothetical prote	654	14	53.8	295	2	G65018	ethanolamine ammon
582	14	53.8	249	2	T21785	hypothetical prote	655	14	53.8	295	2	B85887	ethanolamine ammon
583	14	53.8	250	2	T31836	hypothetical prote	656	14	53.8	295	2	G91042	ethanolamine ammon
584	14	53.8	250	2	T31835	hypothetical prote	657	14	53.8	295	2	C90740	probable transmemb
585	14	53.8	251	2	T14407	pectin esterase ho	658	14	53.8	295	2	A46017	zinc finger protei
586	14	53.8	251	2	A55523	hypothetical prote	659	14	53.8	295	2	B85590	probable transmemb
587	14	53.8	251	2	B39685	GRSAG protein 2.1	660	14	53.8	295	2	B64818	probable membrane
588	14	53.8	251	2	AB0213	conserved hypochet	661	14	53.8	296	2	H87511	hypothetical prote
589	14	53.8	252	2	S61178	hypothetical prote	662	14	53.8	297	2	T18660	hypothetical prote
590	14	53.8	253	2	T04059	hypothetical prote	663	14	53.8	297	2	A98294	hypothetical prote
591	14	53.8	253	2	B89130	protein F52E1.2 [i	664	14	53.8	297	2	AH2989	permease [imported
592	14	53.8	254	2	AB1304	dihydroorotate deh	665	14	53.8	297	2	S06267	surface antigen H
593	14	53.8	254	2	AB1676	sugar fermentation	666	14	53.8	298	2	AF0813	ethanolamine ammon
594	14	53.8	256	2	F82304	hypothetical prote	667	14	53.8	299	2	T43989	minor capsid prote
595	14	53.8	257	2	T28946	hypothetical prote	668	14	53.8	300	2	C88970	transcription regu
596	14	53.8	257	2	T47491	hypothetical prote	669	14	53.8	301	2	F82287	transcription regu
597	14	53.8	258	2	B82349	uridine phosphoryl	670	14	53.8	301	2	B84325	tRNA-pseudouridine
598	14	53.8	259	1	IOH01	inulin-like grow	671	14	53.8	302	2	D65054	hypothetical prote
599	14	53.8	259	1	T27783	hypothetical prote	672	14	53.8	302	2	H84731	hypothetical prote
600	14	53.8	259	2	B84427	hypothetical prote	673	14	53.8	302	2	UC2071	chitinase (EC 3.2.
601	14	53.8	262	2	T05084	hypothetical prote	674	14	53.8	304	2	A72596	hypothetical prote
602	14	53.8	263	2	S44668	ZK370.7 protein -	675	14	53.8	305	2	T28005	hypothetical prote
603	14	53.8	264	2	T09324	capsid protein - h	676	14	53.8	306	2	S32834	methylviologen-red
604	14	53.8	264	2	AE2274	hypothetical prote	677	14	53.8	310	2	T41985	hypothetical prote
605	14	53.8	265	2	S62363	D-2-chlorotriptoni	678	14	53.8	311	2	A28446	transferrin - mous
606	14	53.8	267	2	C81651	conserved hypochet	679	14	53.8	311	2	UC7873	L-thiamose-binding
607	14	53.8	267	2	H86548	polymorphic outer	680	14	53.8	312	2	T09443	zinc metalloprotei
608	14	53.8	270	2	I37278	complement factor	681	14	53.8	312	2	B82276	hypothetical prote
609	14	53.8	270	2	C96731	unknown protein F5	682	14	53.8	312	2	T17969	hypothetical prote
610	14	53.8	270	2	P89632	protein F13E6.3 [i	683	14	53.8	312	2	T45061	hypothetical prote
611	14	53.8	271	2	T40238	hypothetical prote	684	14	53.8	314	2	T28879	hypothetical prote
612	14	53.8	272	2	T15351	hypothetical prote	685	14	53.8	317	2	A33985	wound-inducible ch
613	14	53.8	272	2	T21789	hypothetical prote	686	14	53.8	318	2	T04403	probable chitinase

687	14	53.8	319	2	H83253	conserved hypothet	760	14	53.8	369	2	S60186	peridin-chloroph
688	14	53.8	319	2	B86395	probable origin re	761	14	53.8	370	2	S15013	wnt-1 protein - ze
689	14	53.8	320	2	S3670	chitinase (EC 3.2.	762	14	53.8	372	1	W2WJ21	E2 protein - human
690	14	53.8	320	2	T28992	hypothetical prote	763	14	53.8	372	2	T31060	hypothetical prote
691	14	53.8	322	1	A29328	membrane-associate	764	14	53.8	372	2	T45410	hypothetical prote
692	14	53.8	322	2	H87131	probable transcrip	765	14	53.8	372	2	T29359	hypothetical prote
693	14	53.8	323	2	G87358	diene lactone hydro	766	14	53.8	373	2	A70856	probable 1ppz prot
694	14	53.8	323	2	T33279	hypothetical prote	767	14	53.8	373	2	T47488	hypothetical prote
695	14	53.8	323	2	A47172	transforming growt	768	14	53.8	375	2	G03300	pancreatic polypep
696	14	53.8	324	2	JC2395	Fas antigen precu	769	14	53.8	375	2	T31812	neuropeptide Y/pep
697	14	53.8	325	2	I65354	stem cell antigen	770	14	53.8	375	2	S63685	neuropeptide Y rec
698	14	53.8	327	2	F96814	hypothetical prote	771	14	53.8	375	2	S52143	amac protein - Erw
699	14	53.8	328	2	G89152	protein C24B5.5 f1	772	14	53.8	376	2	A96720	hypothetical prote
700	14	53.8	329	2	T22636	hypothetical prote	773	14	53.8	378	2	B64044	rod shape-determ
701	14	53.8	329	2	T18619	hypothetical prote	774	14	53.8	379	2	G82380	hypothetical prote
702	14	53.8	331	2	T06227	peroxidase (EC 1.1	775	14	53.8	380	2	C86754	Similar to part of
703	14	53.8	331	2	C71468	probable glycerol-	776	14	53.8	381	2	G83835	sulfate adenyllyl r
704	14	53.8	331	2	B83486	cobalamin biosynth	777	14	53.8	382	2	T27076	hypothetical prote
705	14	53.8	332	2	T04484	probable chitinase	778	14	53.8	382	2	B46133	neuropeptide Y/pep
706	14	53.8	332	2	T21458	hypothetical prote	779	14	53.8	382	2	S27388	neuropeptide Y rec
707	14	53.8	332	2	A36663	nautilus protein -	780	14	53.8	382	2	I59565	stem cell antigen
708	14	53.8	333	2	D88970	protein F15E1.11	781	14	53.8	382	2	E91083	hypothetical membr
709	14	53.8	333	2	T34498	hypothetical prote	782	14	53.8	382	2	P85928	hypothetical prote
710	14	53.8	334	2	B81272	probable sugar nuc	783	14	53.8	383	1	VGBBKG	glycoprotein precu
711	14	53.8	336	2	T15727	hypothetical prote	784	14	53.8	384	2	A45490	neuropeptide Y/pep
712	14	53.8	340	2	JC7695	G protein-coupled	785	14	53.8	384	2	S64735	retrovirus-related
713	14	53.8	341	1	WNVZHI	antithrombin-f11 h	786	14	53.8	384	2	T38544	probable exopolyp
714	14	53.8	342	1	H89980	conserved hypothet	787	14	53.8	385	2	S68780	dopamine D1-like r
715	14	53.8	344	2	E72173	D2R protein - vari	788	14	53.8	385	2	S53718	homeotic protein d
716	14	53.8	344	2	S46871	B1R protein - var	789	14	53.8	385	2	A54785	preadipocyte facto
717	14	53.8	344	2	T28611	hypothetical prote	790	14	53.8	386	2	S72168	dopamine receptor
718	14	53.8	344	2	A70771	probable hydrolase	791	14	53.8	386	2	T12527	hypothetical prote
719	14	53.8	345	1	WNVZM2	antithrombin-f11 h	792	14	53.8	386	2	T49935	hypothetical prote
720	14	53.8	345	1	UN0465	apolipoprotein H p	793	14	53.8	387	2	B71611	hypothetical prote
721	14	53.8	345	2	T32018	hypothetical prote	794	14	53.8	388	2	S51316	prostaglandin E re
722	14	53.8	346	2	AF1820	siatoglycoproteina	795	14	53.8	388	2	I38750	prostaglandin recep
723	14	53.8	346	2	T45748	hypothetical prote	796	14	53.8	390	2	S56560	hypothetical 43. 6K
724	14	53.8	346	2	T47411	hypothetical prote	797	14	53.8	390	2	B91291	hypothetical prote
725	14	53.8	347	2	S60428	hypothetical prote	798	14	53.8	390	2	D66132	hypothetical prote
726	14	53.8	347	2	UQ2359	wheat aluminum ind	799	14	53.8	392	2	A60777	keratin 2, type I,
727	14	53.8	348	2	S64041	probable membrane	800	14	53.8	392	2	T33505	hypothetical prote
728	14	53.8	348	2	S75548	siatoglycoproteina	801	14	53.8	395	2	S58157	hypothetical prote
729	14	53.8	349	2	S19873	single-stranded DN	802	14	53.8	396	1	A42919	mevalonate kinase
730	14	53.8	349	2	AF3401	enoyl-CoA hydratase	803	14	53.8	397	2	T08345	hypothetical prote
731	14	53.8	349	2	S12863	G protein-coupled	804	14	53.8	398	2	AH0482	conserved hypothet
732	14	53.8	352	2	C96643	hypothetical prote	805	14	53.8	398	2	T19421	hypothetical prote
733	14	53.8	352	2	S77448	hypothetical prote	806	14	53.8	400	2	S32804	beta-3-adrenergic
734	14	53.8	354	2	T26699	hypothetical prote	807	14	53.8	400	2	A41679	beta-3-adrenergic
735	14	53.8	354	2	G86276	hypothetical prote	808	14	53.8	400	2	A53281	beta-3-adrenergic
736	14	53.8	356	2	A25918	thrombomodulin - b	809	14	53.8	400	2	D89810	conserved hypothet
737	14	53.8	356	2	T18590	hypothetical prote	810	14	53.8	401	2	E96570	hypothetical prote
738	14	53.8	357	2	A71295	probable alanine r	811	14	53.8	402	2	T39591	cyclin - fission Y
739	14	53.8	357	2	A97697	hypothetical prote	812	14	53.8	403	2	C84228	probable trypsin Y
740	14	53.8	358	2	D84901	hypothetical prote	813	14	53.8	403	2	A31847	hypothetical prote
741	14	53.8	358	2	T23802	hypothetical prote	814	14	53.8	403	2	T49480	hypothetical prote
742	14	53.8	359	2	AD3520	galactoside transp	815	14	53.8	405	2	S65459	beta-3-adrenergic
743	14	53.8	361	2	T25305	hypothetical prote	816	14	53.8	405	2	B36340	alpha(1, 3)-fucosyl
744	14	53.8	361	2	C65132	hypothetical 38.6	817	14	53.8	406	2	A43850	glutamate N-acetyl
745	14	53.8	362	1	LPBCRA	rate lipoprotein A	818	14	53.8	406	2	C81017	amino-acid N-acety
746	14	53.8	362	2	C85563	a minor lipoprotei	819	14	53.8	407	2	T19895	hypothetical prote
747	14	53.8	362	2	G90712	hypothetical prote	820	14	53.8	408	1	QRHUB8	beta-3-adrenergic
748	14	53.8	362	2	T20222	dopamine D1 recept	821	14	53.8	408	1	QRHUB8	immediate-early pr
749	14	53.8	363	2	I50475	triacylglycerol f1	822	14	53.8	408	2	B83333	probable MFS trans
750	14	53.8	363	2	A34959	probable acyl-CoA	823	14	53.8	408	2	A72869	early gene transac
751	14	53.8	363	2	A81134	acyl-CoA dehydroge	824	14	53.8	410	2	B84144	hypothetical prote
752	14	53.8	363	2	C81888	probable acyl-CoA	825	14	53.8	412	2	B81176	hypothetical prote
753	14	53.8	364	2	JC2115	prostaglandin E re	826	14	53.8	412	2	T26056	beta-2-adrenergic
754	14	53.8	364	2	S65009	collagen alpha 2(I	827	14	53.8	413	1	QRHUB2	beta-2-adrenergic
755	14	53.8	365	2	S10847	neuropeptide Y/pep	828	14	53.8	413	2	UC2520	extracellular sucr
756	14	53.8	366	1	S71152	GTP-binding protei	829	14	53.8	414	1	QRHUB3	beta-fructofuranos
757	14	53.8	367	2	I51426	pectate lyase (EC	830	14	53.8	414	1	I53040	beta-3-adrenergic
758	14	53.8	367	2	S15716	cinamoyl CoA redu	831	14	53.8	415	2	I53040	beta-2 adrenergic
759	14	53.8	368	2	T46643		832	14	53.8	415	2	A64299	coenzyme F420 hydr

833	14	53.8	415	2	S37690	DNA-binding protei	906	14	53.8	456	2	C71402	probable glucosylt
834	14	53.8	415	2	S34140	DNA-binding protei	907	14	53.8	457	2	I51660	dopamine D1B recep
835	14	53.8	415	2	T26057	hypothetical prote	908	14	53.8	457	2	T32244	hypothetical prote
836	14	53.8	416	2	T45051	hypothetical prote	909	14	53.8	459	2	A25928	cellulase (EC 3.2.
837	14	53.8	418	1	QRHYB2	beta-2-adrenergic	910	14	53.8	459	2	A56849	dopamine recep-tor
838	14	53.8	418	2	G02953	beta-3-adrenergic	911	14	53.8	460	2	T39462	argininosuccinate
839	14	53.8	418	2	S00260	beta-2-adrenergic	912	14	53.8	460	2	T48137	copper amine oxida
840	14	53.8	418	2	S10855	beta-2-adrenergic	913	14	53.8	461	2	H64087	L-tryptophanase se
841	14	53.8	419	2	S14421	GTP-binding regu	914	14	53.8	462	2	A10492	L-tryptophanase se
842	14	53.8	420	2	S06579	finger protein (c1	915	14	53.8	463	2	A48375	cellulase (EC 3.2.
843	14	53.8	420	2	A53531	oncofetal trophob	916	14	53.8	463	2	B56849	dopamine recep-tor
844	14	53.8	421	2	T51055	hypothetical prote	917	14	53.8	464	2	UC7143	endoglucanase I -
845	14	53.8	423	2	A49772	cuticle protein cu	918	14	53.8	464	2	S12591	beta-1-adrenergic
846	14	53.8	424	2	T20032	hypothetical prote	919	14	53.8	464	2	T34332	hypothetical prote
847	14	53.8	424	2	T01383	GTPase-activating	920	14	53.8	465	2	I51661	dopamine D1C recep
848	14	53.8	424	2	T33552	hypothetical prote	921	14	53.8	466	2	S36794	beta-1-adrenergic
849	14	53.8	426	2	D88103	protein M10G1.6 f	922	14	53.8	467	2	T31522	hypothetical prote
850	14	53.8	427	2	S17148	alpha-chromidin rec	923	14	53.8	468	2	S02514	n1B protein - Kie
851	14	53.8	428	2	A55044	beta-4C-adrenergic	924	14	53.8	468	2	T22598	hypothetical prote
852	14	53.8	428	2	S03677	finger protein (c1	925	14	53.8	468	2	C95405	probable sulfite o
853	14	53.8	428	2	T48008	hypothetical prote	926	14	53.8	470	2	S26841	retrovirus-related
854	14	53.8	429	2	S09141	Ndt4 intron 2 prot	927	14	53.8	470	2	AF2828	FAD dependent oxid
855	14	53.8	430	2	S58169	porphobilinogen sy	928	14	53.8	470	2	T49175	hypothetical prote
856	14	53.8	430	2	T16079	hypothetical prote	929	14	53.8	470	2	T45816	hypothetical prote
857	14	53.8	430	2	E90869	hypothetical prote	930	14	53.8	471	2	S18446	variant surface gl
858	14	53.8	430	2	T25962	hypothetical prote	931	14	53.8	471	2	S76290	hypothetical prote
859	14	53.8	430	2	D85749	hypothetical prote	932	14	53.8	472	2	S72909	uroporphyrin-III C
860	14	53.8	430	2	H64883	conserved hypothet	933	14	53.8	474	2	S50957	probable membrane
861	14	53.8	431	2	T18753	hypothetical prote	934	14	53.8	475	2	F70972	probable amidase -
862	14	53.8	431	2	G72285	hybrid cluster 14f	935	14	53.8	475	2	A41271	dopamine receptor
863	14	53.8	431	2	F85097	hypothetical prote	936	14	53.8	477	1	DYH0D5	dopamine receptor
864	14	53.8	432	2	A43448	thrombin receptor	937	14	53.8	477	1	QRHUB1	beta-1-adrenergic
865	14	53.8	433	2	T21232	hypothetical prote	938	14	53.8	477	2	S71400	estrogen receptor
866	14	53.8	433	2	F66974	aspartyl aminopept	939	14	53.8	477	2	D97606	glycolate oxidase
867	14	53.8	434	2	G97610	hypothetical metab	940	14	53.8	478	2	G84576	probable beta-1,3-
868	14	53.8	434	2	AC2833	hypothetical prote	941	14	53.8	480	2	A63491	3-phosphoshikimate
869	14	53.8	434	2	S58476	non-replicative mo	942	14	53.8	480	2	I53053	beta 1 adrenergic
870	14	53.8	434	2	T47575	hypothetical prote	943	14	53.8	480	2	T00971	probable disease r
871	14	53.8	434	2	A38282	p58 galactosyltran	944	14	53.8	480	2	T20168	hypothetical prote
872	14	53.8	435	2	T16470	hypothetical prote	945	14	53.8	480	2	T23608	hypothetical prote
873	14	53.8	435	2	F86601	protoporphyrinogen	946	14	53.8	482	2	UC5092	beta-adrenergic re
874	14	53.8	435	2	D72023	protoporphyrinogen	947	14	53.8	483	2	A25896	retrovirus-related
875	14	53.8	436	2	A81518	hypothetical prote	948	14	53.8	483	2	S52567	GPI-anchor biosynt
876	14	53.8	436	2	AB2417	hypothetical prote	949	14	53.8	484	1	A46217	phosphoprotein pho
877	14	53.8	437	2	S34959	NAH2 dehydrogenas	950	14	53.8	484	2	S36339	phorboloxin II - V
878	14	53.8	437	2	T48035	hypothetical prote	951	14	53.8	484	2	T10668	hypothetical prote
879	14	53.8	439	2	A36385	surface antigen se	952	14	53.8	485	1	A55731	GPI-anchor biosynt
880	14	53.8	439	2	T09214	choline monooxygen	953	14	53.8	485	2	I52484	gene Pig-A protein
881	14	53.8	440	2	B71858	adenylosuccinate 1	954	14	53.8	485	2	T49237	hypothetical prote
882	14	53.8	440	2	H64658	adenylosuccinate 1	955	14	53.8	486	1	JC7241	phosphoprotein pho
883	14	53.8	440	2	T32190	hypothetical prote	956	14	53.8	486	1	COBYW2	WH12 protein - Yea
884	14	53.8	442	2	T10718	anthranilate N-ben	957	14	53.8	486	2	B55886	dopamine receptor
885	14	53.8	442	2	C86471	51.1K hypothetical	958	14	53.8	487	1	DYRTD1	dopamine receptor
886	14	53.8	444	2	C55886	dopamine receptor	959	14	53.8	487	2	T49424	hypothetical prote
887	14	53.8	444	2	T20374	hypothetical prote	960	14	53.8	487	2	T05127	hypothetical prote
888	14	53.8	445	2	T10717	anthranilate N-ben	961	14	53.8	488	1	URBYM	funarate hydratase
889	14	53.8	445	2	T10719	anthranilate N-ben	962	14	53.8	489	2	A49388	catalase (EC 1.11.
890	14	53.8	446	1	DYH0D1	dopamine receptor	963	14	53.8	489	2	D87551	glutaryl-L-lysine (Gln)
891	14	53.8	446	2	T47217	dopamine receptor	964	14	53.8	491	2	UC2497	cyclin E - mouse
892	14	53.8	446	2	PC4393	IIV2 protein - OVI	965	14	53.8	493	2	B77724	glutaryl-L-lysine amid
893	14	53.8	446	2	T10711	anthranilate N-ben	966	14	53.8	493	2	B71725	glutaryl-L-lysine amid
894	14	53.8	446	2	B70543	hypothetical prote	967	14	53.8	493	2	T25357	hypothetical prote
895	14	53.8	446	2	T14542	choline monooxygen	968	14	53.8	496	2	S37379	catalase (EC 1.11.
896	14	53.8	447	2	S25817	hypothetical prote	969	14	53.8	497	2	S47890	exoenzyme 1 - fr
897	14	53.8	448	2	B85077	hypothetical prote	970	14	53.8	498	2	D86298	hypothetical prote
898	14	53.8	449	2	E36676	hypothetical prote	971	14	53.8	499	2	S63465	dihydrofolipamide d
899	14	53.8	450	2	F82659	UDP-glucose dehydr	972	14	53.8	500	1	EJ0258	cyclochrome P450 2D
900	14	53.8	450	2	A55886	dopamine receptor	973	14	53.8	504	2	JQ2151	macron-type locus
901	14	53.8	450	2	T47450	n-acetylglucosamin	974	14	53.8	506	2	A40218	envelop glycoprote
902	14	53.8	450	2	G66670	hypothetical prote	975	14	53.8	506	2	T01292	hypothetical prote
903	14	53.8	451	2	I51659	dopamine D1A recep	976	14	53.8	508	2	D85056	probable aspartic
904	14	53.8	453	2	G86396	protein TN9.2 (1m	977	14	53.8	509	1	VCFPGI	retrovirus-related
905	14	53.8	453	2	T09745	myd-related protei	978	14	53.8	509	2	T00627	hypothetical prote

979 14 53.8 510 1 S15620 l1 protein - human
980 14 53.8 510 1 S15627 l1 protein - human
981 14 53.8 511 2 S44275 dopamine receptor
982 14 53.8 511 2 UC7682 spermatogenesis as
983 14 53.8 513 2 T38866 conserved hypothet
984 14 53.8 515 2 D73119 probable glucose-6
985 14 53.8 516 2 T33269 hypothetical prote
986 14 53.8 516 2 S40996 hypothetical prote
987 14 53.8 517 2 A71707 l1polysaccharide
988 14 53.8 519 2 S78089 G-protein signalin
989 14 53.8 520 2 T23545 hypothetical prote
990 14 53.8 522 2 S71821 probable interleuk
991 14 53.8 523 1 S61713 carboxypeptidase C
992 14 53.8 524 1 P1M158 l1 protein - human
993 14 53.8 527 1 SAHUP saposin precursor
994 14 53.8 528 2 B88989 protein F02C9.2 l1
995 14 53.8 528 2 T33527 hypothetical prote
996 14 53.8 529 2 S46116 probable regulator
997 14 53.8 530 2 JCS939 estrogen receptor
998 14 53.8 531 2 T32276 hypothetical prote
999 14 53.8 532 2 S18448 variant surface g1
1000 14 53.8 533 2 T31002 hypothetical prote

ALIGNMENTS

RESULT 1
S59448
hypothetical protein YMR206w - Yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YMR325.07
C:Species: Saccharomyces cerevisiae
C>Date: 30-Nov-1995 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C:Accession: S59448
R:Odell, C.; Bowman, S.
submitted to the EMBL Data Library, March 1995
A:Reference number: S59441
A:Accession: S59448
A:Molecule type: DNA
A:Residues: 1-313 <ODE>
A:Cross-references: UNIPROT:Q03695; UNIPARC:UPI000013B90A; EMBL:Z48755; NID:G736296; PII
A:Experimental source: strain AB972
C:Genetics:
A:Gene: MIPS:YMR206w
A:Cross-references: SGD:S0004819
A:Map position: 13R
Query Match 69.2%; Score 18; DB 2; Length 313;
Best Local Similarity 20.0%; Pred. No. 0.44; Mismatches 8; Indels 0; Gaps 0;
Matches 2; Conservative 0;
QY 1 CXXXXXXXXXC 10
DB 245 CXXXXXXXXXC 254
RESULT 2
E75433
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: E75433
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uitterlind, T.; Zaleski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: E75433
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-480 <WHI>
A:Cross-references: UNIPROT:Q9R867; UNIPARC:UPI00000C18A8; GB:AE001963; GB:AE000513; NID

A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1142
A:Map position: 1
C:Superfamily: Deinococcus radiodurans hypothetical protein DR1142

Query Match 69.2%; Score 18; DB 2; Length 480;
Best Local Similarity 20.0%; Pred. No. 0.51; Mismatches 8; Indels 0; Gaps 0;
Matches 2; Conservative 0;

QY 1 CXXXXXXXXXC 10
DB 464 CASAAATTAAC 473

RESULT 3
JQ1560
hypothetical 20.6K protein - Lymantria dispar nuclear polyhedrosis virus
N:Alternate names: hypothetical protein 4
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdNPV
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: JQ1560
R:Bjornson, R.M.; Rohmann, G.F.
J. Gen. Virol. 73, 1499-1504, 1992
A:Title: Nucleotide sequence of the polyhedron envelope protein gene region of the Lyman
A:Reference number: PQ0339; MUID:92300345; PMID:1607868
A:Accession: JQ1560
A:Molecule type: DNA
A:Residues: 1-194 <BJO>
A:Cross-references: UNIPROT:P36868; UNIPARC:UPI000017A7BF; DBJ:DI0836

Query Match 65.4%; Score 17; DB 2; Length 194;
Best Local Similarity 20.0%; Pred. No. 1.6; Mismatches 8; Indels 0; Gaps 0;
Matches 2; Conservative 0;

QY 1 CXXXXXXXXXC 10
DB 74 CSAAMTSSSC 83

RESULT 4
JC4082
coat protein - Cymbidium mosaic virus
C:Species: Cymbidium mosaic virus
C>Date: 21-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: JC4082
R:Ryu, K.H.; Yoon, K.E.; Park, W.M.
Gene 156, 303-304, 1995
A:Title: Cloning and sequencing of a cDNA encoding the coat protein of a Korean isolate
A:Reference number: JC4082; MUID:95278762; PMID:7758973
A:Accession: JC4082
A:Molecule type: mRNA
A:Residues: 1-220 <RYU>
A:Cross-references: UNIPROT:Q66152; UNIPARC:UPI00000F22DE; EMBL:X81051; NID:G897718; PII
A:Note: The authors translated the codon GTG for residue 161 as Leu
C:Superfamily: potato virus coat protein
C:Keywords: coat protein

Query Match 65.4%; Score 17; DB 2; Length 220;
Best Local Similarity 20.0%; Pred. No. 1.7; Mismatches 8; Indels 0; Gaps 0;
Matches 2; Conservative 0;

QY 1 CXXXXXXXXXC 10
DB 166 CSAATLTATC 175

RESULT 5
AB2920
cobalamin biosynthetic protein [imported] - Agrobacterium tumefaciens (strain C58, Dupon
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AB2920

R.Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, H.; Karp, P.; Gillet, W.; Grant, C.; Guenther, D.; Kuyavlin, T.; Levy, R.; Li, M.; McClellan, S. Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.

A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AB2920

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-376 <KUR>

A:Cross-references: UNIPROT:Q8UB06; UNIPARC:UPI0000164772; GB:AE008688; PIDN:AAL43776.1;

A:Experimental source: strain C58 (Dugont)

A:Genetics:

A:Gene: chid

A:Map position: circular chromosome

C:Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 65.4%; Score 17; DB 2; Length 376;

Best Local Similarity 20.0%; Pred. No. 2;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10

DB 17 CAAATKAC 26

RESULT 6

C97694

cobalamin biosynthetic protein chid (PA2908) [imported] - *Agrobacterium tumefaciens* (str C:Species: *Agrobacterium tumefaciens*

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004

C:Accession: C97694

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Woliam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens* C97694

A:Accession: C97694

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-388 <KUR>

A:Cross-references: UNIPROT:Q8UB06; UNIPARC:UPI00000D2014; GB:AE007869; PIDN:AAK86508.1;

C:Genetics:

A:Gene: AGR_C_5073

A:Map position: circular chromosome

C:Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 65.4%; Score 17; DB 2; Length 388;

Best Local Similarity 20.0%; Pred. No. 2;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10

DB 29 CAAATKAC 38

RESULT 7

T23167

hypothetical protein K01C8.2 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T23167

R:Time, M.

submitted to the EMBL Data Library, April 1995

A:Reference number: Z19702

A:Accession: T23167

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-389 <WIL>

A:Cross-references: UNIPROT:Q21081; UNIPARC:UPI0000080155; EMBL:Z49068; PIDN:CAA8865.1;

A:Experimental source: clone K01C8

C:Genetics:

A:Gene: CBSP:K01C8.2

A:Map position: 2

A:introns: 54/2; 146/3; 208/3; 283/1; 379/3

Query Match 65.4%; Score 17; DB 2; Length 389;

Best Local Similarity 20.0%; Pred. No. 2;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10

DB 123 CSSSISTSSC 132

RESULT 8

T37314

probable kexin (EC 3.4.21.61) - *Caenorhabditis elegans* (fragment)

N:Alternate names: b1stcrase 4

C:Species: *Caenorhabditis elegans*

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000

C:Accession: T37314

R:Thacker, C.; Peters, K.; Strayko, M.; Rose, A.M.

Genes Dev. 9, 956-971, 1995

A:Title: The b1-4 locus of *Caenorhabditis elegans* encodes structurally distinct kex2/su

A:Reference number: Z21679; MUID:95293228; PMID:7774813

A:Accession: T37314

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-570 <THA>

A:Cross-references: UNIPARC:UPI00001688E1; EMBL:L29440; NID:g459702; PIDN:AAA98752.1; PI

C:Genetics:

A:Gene: b1-4

A:Map position: 1

C:Keywords: alternative splicing; hydrolase; serine proteinase

Query Match 65.4%; Score 17; DB 2; Length 570;

Best Local Similarity 20.0%; Pred. No. 2.3;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10

DB 311 CTSSSATSC 320

RESULT 9

S33575

dnak-type molecular chaperone precursor, mitochondrial - *Leishmania major*

N:Alternate names: heat shock protein 70-related protein; mitochondrial stress protein

C:Species: *Leishmania major*

C:Date: 09-Dec-1997 #sequence_revision 09-Dec-1997 #text_change 31-Dec-2004

C:Accession: S33575; S78090; S05438

R:Searle, S.; McCrossan, M.V.; Smith, D.F.

J. Cell Sci. 104, 1091-1100, 1993

A:Title: Expression of a mitochondrial stress protein in the protozoan parasite *Leishman*

A:Reference number: S33575; MUID:93300981; PMID:8314893

A:Accession: S33575

A:Molecule type: DNA

A:Residues: 1-634 <SEA>

A:Cross-references: UNIPROT:P12076; UNIPARC:UPI0000177D24; EMBL:X64137

R:Smith, D.F.

submitted to the EMBL Data Library, January 1992

A:Reference number: S78090

A:Accession: S78090

A:Molecule type: DNA

A:Residues: 1-460,481-500, 'QGERIRASNGIRGEFDSG', 501-634 <SMT>

A:Cross-references: UNIPARC:UPI000012CC5C; EMBL:X64137; NID:g911289; PIDN:CAA45498.1; PI

R:Searle, S.; Campos, A.J.R.; Coulson, R.M.R.; Spithill, T.W.; Smith, D.F.

Nucleic Acids Res. 17, 5081-5095, 1989

A:Title: A family of heat shock protein 70-related genes are expressed in the promastigote

A:Reference number: S05438; MUID:89345072; PMID:2762121

A:Accession: S05438

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-249 <SEP>
A:Cross-references: UNIPARC:UPI000016BPLC, EMBL:X14574, NID:g9561, PIDN:CAA32713.1; PID:
C:Genetics:
A:Gene: Lmhsp70.1
A:Genome: nuclear
C:Function:
A:Description: involved in protein folding and assembling/disassembling of protein comp
C:Superfamily: bcr protein
C:Keywords: ATP; mitochondrion; molecular chaperone; stress-induced protein
F:1-23/Domain: transit peptide (mitochondrion) #status predicted <TMP>
F:24-634/Product: heat shock protein 70, mitochondrial #status predicted <MAT>

Query Match 65.4%; Score 17; DB 2; Length 634;
Best Local Similarity 20.0%; Pred. No. 2.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 7 CGSAAASAC 16

RESULT 10
D87803
protein hli-4D [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 21-Jun-2002
C:Accession: D87803
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elg
A:Accession: D87803
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-942 <STO>
A:Cross-references: UNIPARC:UPI000016B639; GB:chr_I; PIDN:AAB96754.1; PID:g2773243; GSPT
C:Genetics:
A:Gene: hli-4D
A:Map position: 1
C:Superfamily: kexin; subtilisin homology

Query Match 65.4%; Score 17; DB 2; Length 942;
Best Local Similarity 20.0%; Pred. No. 2.8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 683 CTSSSATSC 692

RESULT 11
S09118
G surface protein 168 - Paramecium primaurelia
C:Species: Paramecium primaurelia
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C:Accession: S09118
R:Prat, A.
J. Mol. Biol. 211, 521-535, 1990
A>Title: Conserved sequences flank variable tandem repeats in two alleles of the G surfa
A:Reference number: S09118; MUID:90172419; PMID:2308165
A:Accession: S09118
A:Molecule type: DNA
A:Residues: 1-2704 <PRA>
A:Cross-references: UNIPROT:PI7053; UNIPARC:UPI000012AE23; EMBL:X52133; NID:g10049; PIDN
C:Genetics:
A:Genetic code: SGC5
C:Superfamily: G surface protein

Query Match 65.4%; Score 17; DB 2; Length 2704;
Best Local Similarity 20.0%; Pred. No. 4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 1632 CVAATPAATTC 1641

RESULT 12
F84312
hypothetical protein Vng1598h [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: F84312
R:Ng, W.V.; Kennedy, S.P.; Mahapatra, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leibauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
A>Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: F84312
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-71 <STO>
A:Cross-references: UNIPROT:Q9HPK1; UNIPARC:UPI0000063920; GB:AE004437; NID:g10581076; P
C:Genetics:
A:Gene: VNG1598H

Query Match 61.5%; Score 16; DB 2; Length 71;
Best Local Similarity 20.0%; Pred. No. 4.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 30 CSRSVSTSTC 39

RESULT 13
E70531
hypothetical protein Rv2706c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: E70531
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: E70531
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-85 <COL>
A:Cross-references: UNIPROT:O07207; UNIPARC:UPI00000C14DC; GB:Z96072; GB:AL123456; NID:g
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv2706c

Query Match 61.5%; Score 16; DB 2; Length 85;
Best Local Similarity 20.0%; Pred. No. 5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 24 CSATVAANVC 33

RESULT 14
S51479
drought-induced protein Di21 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 15-Jul-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S51479; S43176

R.Gosti, F.; Bertauche, N.; Vartanian, N.; Giraudat, J.
Mol. Gen. Genet. 246, 10-18, 1995
A>Title: Abscissic acid-dependent and -independent regulation of gene expression by progl
A:Reference number: S51476; MUID:95124220; PMID:7823304
A:Accession: S51479
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-104 <COS>
A:Cross-references: UNIPROT:Q39084; UNIPARC:UPI000000BF21; EMBL:X78585; NID:9469111; PID
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994
C:Genetics:
A:Gene: D121
C:Superfamily: late embryogenesis-abundant protein leas

Query Match 61.5%; Score 16; DB 2; Length 104;
Best Local Similarity 20.0%; Pred. No. 5.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10
Db 14 CSAAGSLSC 23

RESULT 15
A71249
hypothetical protein PH0248 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C:Accession: A71249
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A>Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: A71249
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-105 <KAW>
A:Cross-references: UNIPROT:Q57986; UNIPARC:UPI0000062D66; GB:AP000001; NID:93236128; PI
A:Experimental source: strain OT3
A>Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH0248

Query Match 61.5%; Score 16; DB 2; Length 105;
Best Local Similarity 20.0%; Pred. No. 5.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10
Db 35 CFTASASLAC 44

RESULT 16
F72549
hypothetical protein APE1682 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: F72549
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: F72549
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <KAW>
A:Cross-references: UNIPROT:Q9YBBS; UNIPARC:UPI000005E048; DDBJ:AF000062; NID:95105244;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1682
C:Superfamily: Aeropyrum pernix hypothetical protein APE1682

Query Match 61.5%; Score 16; DB 2; Length 108;
Best Local Similarity 20.0%; Pred. No. 5.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10
Db 8 CFSASKMAC 17

RESULT 17
A23473
chymotrypsin-like proteinase (EC 3.4.21.-) ~ pig (tentative sequence) (fragments)
N:Alternate names: pancreatic elastase II (misidentification)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 09-Jul-2004
C:Accession: A23473
R:Vered, M.; Gertler, A.; Burstein, Y.
Int. J. Pept. Protein Res. 27, 183-190, 1986
A:Reference number: A23473; MUID:86194934; PMID:3634756
A:Accession: A23473
A:Molecule type: protein
A:Residues: 1-126 <VER>
A:Cross-references: UNIPROT:Q7M325; UNIPARC:UPI0000175C02
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase

Query Match 61.5%; Score 16; DB 2; Length 126;
Best Local Similarity 20.0%; Pred. No. 5.8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10
Db 87 CAGAGASSC 96

RESULT 18
T49498
hypothetical protein B14D6.500 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
C:Accession: T49498
R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49498
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-129 <SCH>
A:Cross-references: UNIPARC:UPI0000179E1D; EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.500
A:Experimental source: BAC clone B14D6, strain OR74A
C:Genetics:
A:Gene: NCSP:B14D6.500
A:Map position: 6
C:Superfamily: Neurospora crassa hypothetical protein B14D6.500

Query Match 61.5%; Score 16; DB 2; Length 129;
Best Local Similarity 20.0%; Pred. No. 5.8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10
Db 90 CSARASISYC 99

RESULT 19
A96746
hypothetical protein T9N14.5 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A96746
R:Rheologis, A.; Becker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luroe, J.S.; Malt, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A66141; MUID:21016719; PMID:11130712
 A:Accession: A96746
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-133 <STO>
 A:Cross-references: UNIPROT:Q9C7T1, UNIPARC:UPI00000450D; GB:AE005173; NID:G10645374; F
 C:Genetics:
 A:Gene: T9N14.5
 A:Map position: 1

Query Match 61.5%; Score 16; DB 2; Length 133;
 Best Local Similarity 20.0%; Pred. No. 5.9;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
 DB 93 CTSSSVSSC 102

RESULT 20
 S09762
 hypothetical protein TRL13 precursor - human cytomegalovirus (strain AD169)
 C:Species: human cytomegalovirus, human herpesvirus 5
 A:Note: host Homo sapiens (man)
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
 C:Accession: S09762
 R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohm, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; M.; Barrall, B.G.
 Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
 A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
 A:Reference number: S09749; MUID:90269039; PMID:2161319
 A:Accession: S09762
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-147 <CHB>
 A:Cross-references: UNIPROT:P16911, UNIPARC:UPI000012D86F; EMBL:X17403; NID:G59591; PIDN
 A:Note: this sequence was submitted to the EMBL Data Library, December 1989
 C:Superfamily: human cytomegalovirus hypothetical protein TRL13
 C:Keywords: glycoprotein
 F.1-25/Domains: signal sequence #status predicted <Sig>
 F.1-147/Product: hypothetical protein TRL13 #status predicted <MAT>
 F.121.35.54,62,99,110,125,133,144/Binding site: carbohydrate (Aam) (covalent) #status pre

Query Match 61.5%; Score 16; DB 2; Length 147;
 Best Local Similarity 20.0%; Pred. No. 6.1;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
 DB 97 CTNTTTTVC 106

RESULT 21
 PNO103
 hypothetical 17K protein - barley stripe mosaic virus
 C:Species: barley stripe mosaic virus, BSMV
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
 C:Accession: PNO103
 R:Kozlov, Y.V.; Afanasiev, B.N.; Rupakov, V.V.; Golova, Y.B.; Kulaeva, O.I.; Dolja, V.V.
 Mol. Biol. (Mosk.) 23, 1080-1090, 1989
 A:Title: The complete nucleotide sequence of barley stripe mosaic virus RNA 3 and its va
 C:Reference number: PNO102; MUID:9006400; PMID:2586501
 A:Accession: PNO103
 A:Molecule type: genomic RNA
 A:Residues: 1-153 <KOZ>

A:Cross-references: UNIPROT:Q07118, UNIPARC:UPI00000F58A5
 C:Genetics:
 A:Map position: segment 3(III)

Query Match 61.5%; Score 16; DB 2; Length 153;
 Best Local Similarity 20.0%; Pred. No. 6.2;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
 DB 10 CGTSTSTVC 19

RESULT 22
 T28088
 hypothetical protein ZK899.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T28088
 R:Kershaw, J.
 submitted to the EMBL Data Library, September 1994
 A:Reference number: Z20468
 A:Accession: T28088
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-161 <WIL>
 A:Cross-references: UNIPROT:Q23659, UNIPARC:UPI000007B9F8, EMBL:Z27140; PIDN:CAA85496.1;
 A:Experimental source: clone ZK899
 C:Genetics:
 A:Gene: CESP:ZK899.1
 A:Map position: X
 A:Introns: 46/3; 89/3; 144/2
 C:Superfamily: *Caenorhabditis elegans* hypothetical protein ZK899.1

Query Match 61.5%; Score 16; DB 2; Length 161;
 Best Local Similarity 20.0%; Pred. No. 6.3;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
 DB 152 CSTSTSTSC 161

RESULT 23
 JQ1252
 hypothetical 16.7K protein - chrysanthemum virus B
 C:Species: chrysanthemum virus B
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
 C:Accession: JQ1252
 R:Levay, K.; Zavriv, S.
 J. Gen. Virol. 72, 2333-2337, 1991
 A:Title: Nucleotide sequence and gene organization of the 3'-terminal region of chrysant
 A:Reference number: JQ1246; MUID:92013948; PMID:1919520
 A:Accession: JQ1252
 A:Molecule type: genomic RNA
 A:Residues: 1-164 <LEV>
 A:Cross-references: UNIPROT:Q7LZW7, UNIPARC:UPI0000179B53

Query Match 61.5%; Score 16; DB 2; Length 164;
 Best Local Similarity 20.0%; Pred. No. 6.3;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
 DB 80 CSLSSAASC 89

RESULT 24
 T34277
 hypothetical protein F46H5.2 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T34277

R.Nhan, M.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F46H5.
A:Reference number: Z21498
A:Accession: T34277
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-229 <RNA>
A:Cross-references: UNIPROT:Q20490; UNIPARC:UPI0000179CAC; EMBL:U41543; PIDN:AB37021.1;
C:Experimental source: strain Bristol NZ; clone F46H5
C:Genetics:
A:Gene: CESP:F46H5.2
A:Map position: X
A:introns: 16/1; 85/2; 151/2

Query Match 61.5%; Score 16; DB 2; Length 229;
Best Local Similarity 20.0%; Pred. No. 7.1;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 23 CEASASNSSC 32

RESULT 25
A45724
pectate lyase (EC 4.2.2.2) - fungus (Fusarium solani)
C:Species: Fusarium solani
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A45724
R:Gonzalez-Candelas, L.; Kolatukudy, P.E.
J. Bacteriol. 174, 6343-6349, 1992
A:Title: Isolation and analysis of a novel inducible pectate lyase gene from the phytora
A:Reference number: A45724; MUID:93015682; PMID:1400187
A:Accession: A45724
A:Status: preliminary
A:Molecule type: DNA; protein
A:Residues: 1-242 <CON>
A:Cross-references: UNIPROT:Q04701; UNIPARC:UPI000006C8B5; GB:M94691; NID:g168155; PIDN:
A:Experimental source: isolate T8
A:Note: sequence extracted from NCBI backbone (NCBIN:115473, NCBIPI:115474)
C:Keywords: carbon-oxygen lyase

Query Match 61.5%; Score 16; DB 2; Length 242;
Best Local Similarity 20.0%; Pred. No. 7.3;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 233 CTVGTSTTSC 242

RESULT 26
S72219
chymotrypsin B - Atlantic cod (fragments)
C:Species: Gadus morhua (Atlantic cod)
C:Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 09-Jul-2004
C:Accession: S72219
R:leeth-Larsen, R.; Asgerisson, B.; Thorolfsson, M.; Norregard-Madsen, M.; Hojrup, P.
Biochim. Biophys. Acta 1297, 49-56, 1996
A:Title: Structure of chymotrypsin variant B from Atlantic cod, Gadus morhua.
A:Reference number: S72219; MUID:96439045; PMID:8841380
A:Accession: S72219
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-14,15-244 <LET>
A:Cross-references: UNIPROT:Q9PM06; UNIPARC:UPI0000175C00; UNIPARC:UPI0000175C01
C:Superfamily: trypsin; trypsin homology
F:15-237/Domain: trypsin homology <TRY>

Query Match 61.5%; Score 16; DB 2; Length 244;
Best Local Similarity 20.0%; Pred. No. 7.3;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 181 CAGAGATSC 190

RESULT 27
A55035
cysteine-rich protein CRP1 - earthworm (Eudryotreus buchholzi)
C:Species: Eudryotreus buchholzi
C:Date: 14-Nov-1994 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: A55035; S45034
R:Willuhn, J.; Schmitt-Wrede, H.P.; Greven, H.; Wunderlich, F.
J. Biol. Chem. 269, 24688-24691, 1994
A:Title: cDNA cloning of a cadmium-inducible mRNA encoding a novel cysteine-rich, non-m
A:Reference number: A55035; MUID:95014230; PMID:7929141
A:Accession: A55035
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-251 <WIL>
A:Cross-references: UNIPROT:Q24774; UNIPARC:UPI000007D243; EMBL:X79344; NID:g488802; PI
C:Superfamily: ultra-high-sulfur keratin

Query Match 61.5%; Score 16; DB 2; Length 251;
Best Local Similarity 20.0%; Pred. No. 7.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 28 CGTSTTAGSC 37

RESULT 28
J01724
EI membrane glycoprotein precursor - canine coronavirus (strain Insavec-1)
N:Alternate names: matrix glycoprotein
C:Species: canine coronavirus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C:Accession: J01724
R:Horsburgh, B.C.; Briarley, I.; Brown, T.D.K.
J. Gen. Virol. 73, 2849-2862, 1992
A:Title: Analysis of a 9.6 kb sequence from the 3' end of canine coronavirus genomic RNA
A:Reference number: P00481; MUID:93057357; PMID:1431811
A:Accession: J01724
A:Molecule type: genomic RNA
A:Residues: 1-262 <HOR>
A:Cross-references: UNIPROT:P36299; UNIPARC:UPI0000138986; DDBJ:D13096; NID:g406193; PI
C:Genetics:
A:Gene: M
C:Superfamily: coronavirus EI membrane glycoprotein
C:Keywords: glycoprotein; matrix protein; transmembrane protein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-262/Product: EI membrane glycoprotein #status predicted <MAT>
F:56-72/Domain: transmembrane #status predicted <TM1>
F:85-101/Domain: transmembrane #status predicted <TM2>
F:115-134/Domain: transmembrane #status predicted <TM3>
F:132-55/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:133-57/Binding site: carbohydrate (Ser) (covalent) #status predicted
F:144/Binding site: carbohydrate (Thr) (covalent) #status predicted

Query Match 61.5%; Score 16; DB 1; Length 262;
Best Local Similarity 20.0%; Pred. No. 7.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 21 CAMTSSTSC 30

RESULT 29
S47537
chymotrypsin (EC 3.4.21.1) precursor - Atlantic cod
C:Species: Gadus morhua (Atlantic cod)

C>Date: 26-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C/Accession: S47537; S43163
R/Gutmundsdottir, A.; Oskarsson, S.; Eakin, A.E.; Craik, C.S.; Bjarnason, J.B.
Biochim. Biophys. Acta 1219, 211-214, 1994
A/Title: Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.
A/Reference number: S47537; MUID:94368660; PMID:8086467
A/Accession: S47537
A/Molecule type: mRNA
A/Residues: 1-263 <GUD>
A/Cross-references: UNIPROT:P47796; UNIPARC:UPI0000128664; EMBL:X78490; NID:G468750; PID
C/Superfamily: trypsin; trypsin homology
C/Keywords: hydrolase; protein digestion; serine proteinase
P:1-18/Domain: signal sequence #status predicted <SIG>
F:19-263/Product: chymotrypsin #status predicted <MAT>
F:34-256/Domain: trypsin homology <TRY>
F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 61.5%; Score 16; DB 2; Length 263;
Best Local Similarity 20.0%; Pred. No. 7.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
DB 200 CAGAGAGASC 209

RESULT 30
A/Accession: A85856
probable elongation factor yelp [imported] - Escherichia coli (strain O157:H7, substrain
C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 05-Oct-2004
C/Accession: A85856
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimmlatta, E.; Potamousis, K.; Apodaca,
Natter 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: A85856
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-275 <STO>
A/Cross-references: UNIPARC:UPI000000288; GB:AE005174; NID:G12516497; PIDN:AA657309.1;
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Gene: yelp
C/Superfamily: translation elongation factor P

Query Match 61.5%; Score 16; DB 2; Length 275;
Best Local Similarity 20.0%; Pred. No. 7.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
DB 54 CTSKSANTTC 63

RESULT 31
A/Accession: B64986
hypothetical 30.9 kD protein in frnB 5' region - Escherichia coli (strain K-12)
C/Species: Escherichia coli
C/Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 05-Oct-2004
C/Accession: B64986
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: B64986
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-275 <BLAT>
A/Cross-references: UNIPARC:UPI0000167DC1; GB:AE000306; GB:U00096; NID:G1788489; PIDN:AA
A/Experimental source: strain K-12, substrain MG1655

C/Genetics:
A/Gene: yelp
C/Superfamily: translation elongation factor P

Query Match 61.5%; Score 16; DB 2; Length 275;
Best Local Similarity 20.0%; Pred. No. 7.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
DB 54 CTSKSANTTC 63

RESULT 32
G91011
probable elongation factor [imported] - Escherichia coli (strain O157:H7, substrain RMD
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 05-Oct-2004
C/Accession: G91011
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: G91011
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-275 <HAY>
A/Cross-references: UNIPARC:UPI000000288; GB:BA000007; PIDN:BA836486.1; PID:G13362532;
A/Experimental source: strain O157:H7, substrain RMD 0509952
C/Genetics:
A/Gene: EC93063
C/Superfamily: translation elongation factor P

Query Match 61.5%; Score 16; DB 2; Length 275;
Best Local Similarity 20.0%; Pred. No. 7.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
DB 54 CTSKSANTTC 63

RESULT 33
T23682
hypothetical protein M02G9.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T23682
R/Matthews, L.
submitted to the EMBL Data Library, November 1996
A/Reference number: Z19781
A/Accession: T23682
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-284 <MTL>
A/Cross-references: UNIPROT:O9XUS0; UNIPARC:UPI0000076A8F; EMBL:Z81573; PIDN:CAB04626.1;
A/Experimental source: clone M02G9
C/Genetics:
A/Gene: CESP:M02G9.3
A/Map position: 2
A/Introns: 20/3; 76/2; 182/3; 223/3

Query Match 61.5%; Score 16; DB 2; Length 294;
Best Local Similarity 20.0%; Pred. No. 7.8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
DB 43 CSSSSSNSTC 52

RESULT 34

C71498
probable yop translocation R - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C/Species: Chlamydia trachomatis
C/Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C/Accession: C71498
R/Stephen, R.S.; Kallman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998
A/Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
A/Reference number: A71570; MWID:99000809; PMID:9784136
A/Accession: C71498
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-306 <RN>
A/Cross-references: UNIPROT:084566; UNIPARC:UPI000003360; GB:AE001327; GB:AE001273; NID: A/Experimental source: serotype D, strain UW-3/Cx
C/Genetics:
A/Genes: yscR

Query Match 61.5%; Score 16; DB 2; Length 306;
Best Local Similarity 20.0%; Pred. No. 7.9;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10
DB 24 CSAAAGASSC 33

RESULT 35

T00500
probable elicitor response element-binding protein WRKY3 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 31-Dec-2004
C/Accession: T00500; C84623
R/Rounsley, S.D.; Lin, X.; Kechum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul submitted to the EMBL Data Library, November 1997
A/Description: Arabidopsis thaliana chromosome II BAC T20D16 genomic sequence.
A/Reference number: Z14164
A/Accession: T00500
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-317 <RN>
A/Cross-references: UNIPROT:022176; UNIPARC:UPI00001387D; EMBL:AC002391; NID:92642427;
A/Experimental source: cultivar Columbia
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MWID:20083487; PMID:10617197
A/Accession: C84623
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-317 <STO>
A/Cross-references: UNIPARC:UPI00001387D; GB:AE002093; NID:92642432; PIDN:AAB87100.1; C/Genetics:
A/Genes: WRKY3; ATSP:T20D16.5; At2g23320
A/Map position: 2
A/Introns: 218/2; 260/2
C/Superfamily: DNA-binding protein WRKY3

Query Match 61.5%; Score 16; DB 2; Length 317;
Best Local Similarity 20.0%; Pred. No. 8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10
DB 203 CASASSGRC 212

RESULT 36

A84792
hypothetical protein At2g37380 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: A84792
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MWID:20083487; PMID:10617197
A/Accession: A84792
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-321 <STO>
A/Cross-references: UNIPROT:09ZUS8; UNIPARC:UPI00000A2519; GB:AE002093; NID:94056491; C/Genetics:
A/Genes: At2g37380
A/Map position: 2
C/Superfamily: Arabidopsis thaliana hypothetical protein At2g39370

Query Match 61.5%; Score 16; DB 2; Length 321;
Best Local Similarity 20.0%; Pred. No. 8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10
DB 306 CSRTSVSTC 315

RESULT 37

J0202
paired-box-containing Pax4 protein - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Jul-2004
C/Accession: J0202
R/Rokuyama, Y.; Yagui, K.; Sakurai, K.; Hashimoto, N.; Saico, Y.; Kanatsuka, A. Biochem. Biophys. Res. Commun. 248, 153-156, 1998
A/Title: Molecular cloning of rat Pax4: Identification of four isoforms in rat insulinoma
A/Reference number: J0202; MWID:98340866; PMID:9675102
A/Accession: J0202
A/Molecule type: mRNA
A/Residues: 1-349 <TOX>
A/Cross-references: UNIPROT:088436; UNIPARC:UPI00001136B; GB:AF053100; NID:93192973; C/Superfamily: paired box transcription factor Pax-4; homeobox homology; paired box homology
C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
F/5-129/Domain: paired box homology <PBH>
F/171-227/Domain: homeobox homology <HOX>

Query Match 61.5%; Score 16; DB 2; Length 349;
Best Local Similarity 20.0%; Pred. No. 8.3;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10
DB 290 CSDDTTSQAC 299

RESULT 38

T04007
hypothetical protein T5L19.160 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C/Accession: T04007
R/Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X submitted to the Protein Sequence Database, March 1999
A/Reference number: Z15184
A/Accession: T04007
A/Molecule type: DNA
A/Residues: 1-380 <BEV>
A/Cross-references: UNIPROT:09T0G1; UNIPARC:UPI00000A0088; EMBL:AL049481
A/Experimental source: cultivar Columbia; BAC clone T5L19
C/Genetics:
A/Map position: 4
A/Introns: 38/3; 83/3; 92/3; 123/1; 137/3; 170/3; 190/1; 207/3; 241/3; 263/3; 301/2; 358/1
A/Note: T5L19.160

```
Query Match          61.5%; Score 16; DB 2; Length 380;
Best Local Similarity 20.0%; Pred. No. 8.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
Db 6 CSTASSSGGC 15

RESULT 39
T26077
hypothetical protein W02A2.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T26077
R:Ainscough, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z20148
A:Accession: T26077
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-397 <MIL>
A:Cross-references: UNIPROT:Q9XUB6; UNIPARC:UPI0000076467; EMBL:Z82286; PIDs: CAB05306.1;
A:Experimental source: clone W02A2
C:Genetics:
A:Gene: CESP.W02A2.3
A:Map position: 4
A:Introns: 15/1; 67/2; 101/3; 259/1; 383/3
C:Superfamily: gliadin

Query Match          61.5%; Score 16; DB 2; Length 397;
Best Local Similarity 20.0%; Pred. No. 8.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
Db 10 CATTAAASC 19

RESULT 40
S41514
RAD52 protein homolog - chicken
C:Species: Gallus gallus (chicken)
C:Date: 25-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 31-Dec-2004
C:Accession: S41514
R:Bezubova, O.Y.; Schmidt, H.; Ostermann, K.; Heyer, W.D.; Buerstedde, J.M.
Nucleic Acids Res. 21, 5945-5949, 1993
A:Title: Identification of a chicken RAD52 homologue suggests conservation of the RAD52
A:Reference number: S41514; MUID:94119695; PMID:8290357
A:Accession: S41514
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-422 <BEZ>
A:Cross-references: UNIPROT:P39022; UNIPARC:UPI0000133034
C:Superfamily: human nuclear cap-binding protein

Query Match          61.5%; Score 16; DB 2; Length 422;
Best Local Similarity 20.0%; Pred. No. 8.8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
Db 17 CTSTNSVAC 26

RESULT 41
C81039
lipopolysaccharide biosynthesis protein, probable NMB1818 [imported] - Neisseria meningi
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: C81039
R:Tetzelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
```

```
Hikey, B.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathavan, J.; Gill, J.; Scariato, V.; Masiyanti, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: C81039
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-473 <TEH>
A:Cross-references: UNIPROT:Q9UXZ7; UNIPARC:UPI0000030867; GB:AE002531; GB:AE002098; NID
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1818

Query Match          61.5%; Score 16; DB 2; Length 473;
Best Local Similarity 20.0%; Pred. No. 9.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
Db 421 CTTSSAAYTC 430

RESULT 42
C81984
probable lipopolysaccharide biosynthesis translocase NMA0643 [imported] - Neisseria meni
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: C81984
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: C81984
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-473 <PAR>
A:Cross-references: UNIPROT:Q9UW9; UNIPARC:UPI0000030868; GB:AL162753; GB:AL157959; NID
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA0641; NMA0643

Query Match          61.5%; Score 16; DB 2; Length 473;
Best Local Similarity 20.0%; Pred. No. 9.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
Db 421 CLASSAAYTC 430

RESULT 43
T36342
probable glutamate decarboxylase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36342
R:Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z21575
A:Accession: T36342
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-475 <OLI>
A:Cross-references: UNIPROT:Q9X8J5; UNIPARC:UPI00000DB08A; EMBL:AL049841; PIDs: CAB42769.
A:Experimental source: strain A3 (2)
C:Genetics:
A:Gene: gad; SCOEDB:SCB9.23
C:Superfamily: Escherichia coli glutamate decarboxylase

Query Match          61.5%; Score 16; DB 2; Length 475;
```


Best Local Similarity 20.0%; Pred. No. 9.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10

DB 131 CSTTGSSEAC 140

RESULT 44

B66460
hypothetical protein F14M2.19 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: B66460

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: B66460

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-486 <STO>

A:Cross-references: UNIPROT:Q9LQ21; UNIPARC:UPI00000A7F31; GB:AE005172; NID:G9665100; PI

A:Map position: 1

C:Genetics:

Query Match 61.5%; Score 16; DB 2; Length 486;

Best Local Similarity 20.0%; Pred. No. 9.3;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10

DB 48 CUSSTGAATC 57

RESULT 45

A55207
glycerol-3-phosphate dehydrogenase g1pd (EC 1.1.-.-) - Pseudomonas aeruginosa

C:Species: Pseudomonas aeruginosa

C:Date: 05-May-1995 #sequence_revision 05-May-1995 #text_change 09-Jul-2004

C:Accession: A55207

R:Schweizer, H.P.; Po, C.

J. Bacteriol. 176, 2184-2193, 1994

A:Title: Cloning and nucleotide sequence of the g1pd gene encoding sn-glycerol-3-phosphat

A:Reference number: A55207; MUID:94209216; PMID:8157588

A:Accession: A55207

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-510 <SCH>

A:Cross-references: UNIPROT:P52111; UNIPARC:UPI000016FCAD; GB:L06231; NID:G450377; PIND:

C:Genetics:

A:Gene: g1pd

C:Superfamily: glycerol-3-phosphate dehydrogenase (aerobic)

C:Keywords: oxidoreductase

Query Match 61.5%; Score 16; DB 2; Length 510;

Best Local Similarity 20.0%; Pred. No. 9.4;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10

DB 194 CTTMSATAAC 203

RESULT 46

G31277
hypothetical protein F52H3.2 - Caenorhabditis elegans

guinate transport protein - Neurospora crassa (tentative sequence)

N:Alternate names: guinate transporter

C:Species: Neurospora crassa

C:Date: 26-Apr-1989 #sequence_revision 26-Apr-1989 #text_change 09-Jul-2004

C:Accession: S04254; G31277

R:Geever, R.F.; Hulet, L.; Baum, J.A.; Tyler, B.M.; Patel, V.B.; Rutledge, B.J.; Case, J.

J. Mol. Biol. 207, 15-34, 1989

A:Title: DNA sequence, organization and regulation of the ga gene cluster of Neurospora

A:Reference number: S04250; MUID:89293848; PMID:2525625

A:Accession: S04254

A:Molecule type: DNA

A:Residues: 1-537 <GE2>

A:Cross-references: UNIPROT:P11636; UNIPARC:UPI0000132E4B; EMBL:X14603; NID:G3060; PIND

C:Genetics:

A:Gene: ga-Y

C:Superfamily: maltose transport protein MAL61

C:Keywords: transmembrane protein

F:22-42/Domain: transmembrane #status predicted <TM01>

F:57-87/Domain: transmembrane #status predicted <TM02>

F:99-119/Domain: transmembrane #status predicted <TM03>

F:132-152/Domain: transmembrane #status predicted <TM04>

F:161-181/Domain: transmembrane #status predicted <TM05>

F:195-215/Domain: transmembrane #status predicted <TM06>

F:286-306/Domain: transmembrane #status predicted <TM07>

F:324-344/Domain: transmembrane #status predicted <TM08>

F:356-376/Domain: transmembrane #status predicted <TM09>

F:390-410/Domain: transmembrane #status predicted <TM10>

F:427-447/Domain: transmembrane #status predicted <TM11>

F:459-479/Domain: transmembrane #status predicted <TM12>

QY 1 CXXXXXXXXXC 10

DB 24 CAAIASFASC 33

RESULT 47

B38418
jockey protein 1 - fruit fly (Drosophila funebris)

C:Species: Drosophila funebris

C:Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 21-Feb-1997

C:Accession: B38418

R:Mizokhi, L.U.; Mazo, A.M.

Proc. Natl. Acad. Sci. U.S.A. 87, 9216-9220, 1990

A:Title: Evidence for horizontal transmission of the mobile element jockey between dipta

A:Reference number: A38418; MUID:91067678; PMID:1701254

A:Accession: B38418

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-580 <MI2>

A:Cross-references: UNIPARC:UPI000017BE6C; GB:M38437

C:Genetics:

A:Gene: FlyBase:Dfunt/jockey

A:Cross-references: FlyBase:FBgn0012310

Query Match 61.5%; Score 16; DB 2; Length 580;

Best Local Similarity 20.0%; Pred. No. 9.9;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10

DB 78 CTVTTSQASC 87

RESULT 48

T22518
hypothetical protein F52H3.2 - Caenorhabditis elegans

hypothetical protein F52H3.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T22518

R.Gardner, A.
 Submitted to the EMBL Data Library, October 1995
 A:Reference number: Z19575
 A:Accession: T22518
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-638 <WIL>
 A:Cross-references: UNIPROT:Q20680; UNIPARC:UPI000013C10D; EMBL:Z66512; PIDD:CAA91322.1;
 A:Experimental source: clone F52H3
 C:Genetics:
 A:Gene: CBSP:F52H3.2
 A:Map position: 2
 A:Introns: 67/1; 157/2; 285/2; 420/1; 566/3
 C:Superfamily: gida protein

Query Match 61.5%; Score 16; DB 2; Length 638;
 Best Local Similarity 20.0%; Pred. No. 10;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 CXXXXXXXXXC 10
 Db 29 CESAAMARC 38

RESULT 49
 C81345
 Hypothetical protein Cj0736 [imported] - Campylobacter jejuni (strain NCTC 11168)
 C:Species: Campylobacter jejuni
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C:Accession: C81345
 R:Parhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillif, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVlier, A.; Whitehead, S.; Barrell, NATURE 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
 A:Reference number: A81250; MUID:20150912; PMID:10688204
 A:Accession: C81345
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-642 <PAR>
 A:Cross-references: UNIPROT:Q9PPH1; UNIPARC:UPI00000C1D12; GB:AL139076; GB:AL111168; NID
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: Cj0736

Query Match 61.5%; Score 16; DB 2; Length 642;
 Best Local Similarity 20.0%; Pred. No. 10;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 CXXXXXXXXXC 10
 Db 375 CSSAAVESC 384

RESULT 50
 T37581
 probable serine-threonine-protein kinase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T37581
 R:Badcock, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
 submitted to the EMBL Data Library, April 1996
 A:Reference number: Z21726
 A:Accession: T37581
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-648 <BAD>
 A:Cross-references: UNIPARC:UPI000017B219; EMBL:Z70721; PIDD:CAA94704.1; GSPDB:GN00066;
 C:Genetics:
 A:Experimental source: strain 972h-, cosmid c12B10
 A:Gene: SPDB:SPAC12B10.14c
 A:Map position: 1

Query Match 61.5%; Score 16; DB 2; Length 648;

Best Local Similarity 20.0%; Pred. No. 10;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 CXXXXXXXXXC 10
 Db 21 COSNASSAC 30

Search completed: January 4, 2006, 16:10:29
 Job time : 12.3478 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 ; Search time 65.3043 Seconds

(without alignments)
108.037 Million cell updates/sec

Title: US-09-932-322-11

Perfect score: 26
Sequence: 1 CXXXXXXXXX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80:*
1: uniprot_sprotl:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	69.2	213	2	Q96KM3 homo sapien
2	18	69.2	263	2	Q586F4 trypanosoma
3	18	69.2	269	2	Q7PRO7 ANOGA
4	18	69.2	281	2	Q5TX10 ANOGA
5	18	69.2	313	1	YME8 YEAST
6	18	69.2	357	2	Q6ZDR4 ORYSA
7	18	69.2	362	2	Q7QJL0 ANOGA
8	18	69.2	480	2	Q9RV87 DEIRA
9	18	69.2	653	1	TMCC1 HUMAN
10	18	69.2	1005	2	Q5H317 XANTHOMAS
11	18	69.2	1061	2	Q4QCK7 LEIIMA
12	18	69.2	52	2	Q8LPD6 HORVU
13	17	65.4	122	2	Q5DCN3 SCHJA
14	17	65.4	128	2	Q96LJ4 HUMAN
15	17	65.4	139	2	Q6ZKM7 ORYSA
16	17	65.4	155	2	Q6H482 ORYSA
17	17	65.4	160	2	Q657M0 ORYSA
18	17	65.4	160	2	Q474H1 TETNG
19	17	65.4	219	2	Q825E8 STRAW
20	17	65.4	220	2	Q6152 GVIIRU
21	17	65.4	256	2	Q6YTS2 ORYSA
22	17	65.4	260	2	Q850Z3 ORYSA
23	17	65.4	304	2	Q5TVE4 ANOGA
24	17	65.4	344	2	Q582B8 TRYYP
25	17	65.4	361	2	Q572Z0 TRYYP
26	17	65.4	364	2	Q4YW87 PLABE
27	17	65.4	368	2	Q7ORD3 GERBY
28	17	65.4	376	1	CBID AGRTS
29	17	65.4	388	2	Q7PY13 ANOGA
30	17	65.4	389	2	Q21081 CAENOT
31	17	65.4	411	2	Q67U23 ORYSA

32	17	65.4	437	2	Q414Q8 GIBBE	Q414Q8 gibberella
33	17	65.4	471	2	Q9YMG7 DROME	Q9YMG7 drosophila
34	17	65.4	475	2	Q7X7A4 ORYSA	Q7X7A4 oryza sativ
35	17	65.4	491	2	Q4P4C7 USTMA	Q4P4C7 ustilla ma
36	17	65.4	520	2	Q4XZV5 PLACH	Q4XZV5 plasmodium
37	17	65.4	556	2	Q803D5 BRABE	Q803D5 brachydanio
38	17	65.4	568	2	Q5GXM1 XANOR	Q5GXM1 xanthomonas
39	17	65.4	626	2	Q4N2Y7 THEPA	Q4N2Y7 thelrella p
40	17	65.4	630	2	Q69TY8 ORYSA	Q69TY8 oryza sativ
41	17	65.4	634	1	HSP71 LEIMA	P12076 leishmania
42	17	65.4	635	2	Q4Q747 LEIMA	Q4Q747 leishmania
43	17	65.4	652	2	Q4Q740 LEIMA	Q4Q740 leishmania
44	17	65.4	660	2	Q4Q744 LEIMA	Q4Q744 leishmania
45	17	65.4	662	2	Q4Q745 LEIMA	Q4Q745 leishmania
46	17	65.4	691	2	Q60E18 ORYSA	Q60E18 oryza sativ
47	17	65.4	699	2	Q819P8 TRYPA	Q819P8 corbicula f
48	17	65.4	705	2	Q516X2 CHLAB	Q516X2 chlamydomon
49	17	65.4	734	2	Q9NKE3 DROME	Q9NKE3 drosophila
50	17	65.4	784	2	Q4YNR2 PLABE	Q4YNR2 plasmodium
51	17	65.4	843	2	Q80GV2 HPEVO	Q80GV2 hepatitis b
52	17	65.4	861	2	Q55NL5 CRYME	Q55NL5 cryptococcu
53	17	65.4	943	1	BLI4 CAEBL	P51559 caenorhabdi
54	17	65.4	1089	2	Q4Q9C2 LEIMA	Q4Q9C2 leishmania
55	17	65.4	1103	2	Q4PDJ1 USTMA	Q4PDJ1 ustilla ma
56	17	65.4	1200	2	Q4RP81 TETNG	Q4RP81 tetradodon n
57	17	65.4	1347	2	Q6CC16 YARLI	Q6CC16 yarowia li
58	17	65.4	1749	2	Q4V217 BURMA	Q4V217 burkholderi
59	17	65.4	1835	2	Q4S914 TETNG	Q4S914 tetradodon n
60	17	65.4	2704	1	G168 PARPR	P17053 parametium
61	17	65.4	2717	1	Q94710 PARTE	Q94710 parametium
62	17	65.4	2721	2	Q76973 PARTE	Q76973 parametium
63	17	65.4	2729	2	Q6P0K6 PARTE	Q6P0K6 parametium
64	17	65.4	68	2	Q6JL92 NEIGO	Q6JL92 neisseria g
65	16	61.5	71	2	Q9HPK1 HALBA	Q9HPK1 halobacteri
66	16	61.5	80	2	Q6RGD0 DROME	Q6RGD0 drosophila
67	16	61.5	85	2	Q07207 MYCTU	Q07207 mycobacteri
68	16	61.5	85	2	Q7YV24 MYCBO	Q7YV24 mycobacteri
69	16	61.5	85	2	Q7NSB4 CHRVO	Q7NSB4 chromobacte
70	16	61.5	82	2	Q5Z713 ORYSA	Q5Z713 oryza sativ
71	16	61.5	98	2	Q5Z2E5 NOCPA	Q5Z2E5 nocardia fa
72	16	61.5	103	2	Q5YF63 TRYPU	Q5YF63 rock bream
73	16	61.5	104	2	Q39084 ARAYH	Q39084 arabidopsis
74	16	61.5	105	2	Q5Y986 PYRHO	Q5Y986 pyrococcus
75	16	61.5	108	2	Q5YBBS ASRBE	Q5YBBS asceropyrum p
76	16	61.5	108	2	Q854Z9 JCAUD	Q854Z9 mycobacteri
77	16	61.5	108	2	Q6Y9M6 LIGQV	Q6Y9M6 ligustrum o
78	16	61.5	110	2	Q7PJV4 ANOGA	Q7PJV4 anopheles g
79	16	61.5	112	2	Q95M57 BOVIN	Q95M57 bos taurus
80	16	61.5	112	2	Q8V6M9 VYIKU	Q8V6M9 halovirus h
81	16	61.5	120	2	Q6ID11 ARAYH	Q6ID11 arabidopsis
82	16	61.5	123	2	Q84WT9 ARAYH	Q84WT9 arabidopsis
83	16	61.5	125	2	Q59Z29 CANAL	Q59Z29 candida alb
84	16	61.5	126	2	Q7M3Z5 PIG	Q7M3Z5 sus scrofa
85	16	61.5	126	2	Q6YUC3 ORYSA	Q6YUC3 oryza sativ
86	16	61.5	126	2	Q8CLV7 YERPE	Q8CLV7 yersinia pe
87	16	61.5	129	2	Q7RUVV NEUPC	Q7RUVV neurospora
88	16	61.5	132	2	Q6UNK8 BOMMO	Q6UNK8 bombyx mori
89	16	61.5	133	2	Q9C7T1 ARAYH	Q9C7T1 arabidopsis
90	16	61.5	137	2	Q9HAI1 HUMAN	Q9HAI1 homo sapien
91	16	61.5	137	2	Q24987 GIALA	Q24987 gliardia lam
92	16	61.5	142	2	Q8MLI9 DROME	Q8MLI9 drosophila
93	16	61.5	142	2	Q08497 HUMAN	Q08497 homo sapien
94	16	61.5	147	1	IRI3 HGWYA	P16811 human cytom
95	16	61.5	147	2	Q6RE18 JNOCA	Q6RE18 rhodococcus
96	16	61.5	149	2	Q4RZ65 TETNG	Q4RZ65 tetradodon n
97	16	61.5	149	2	Q4SZ33 TETNG	Q4SZ33 tetradodon n
98	16	61.5	151	2	Q5H3T3 XANOR	Q5H3T3 xanthomonas
99	16	61.5	153	1	GAMB BSNV	Q08074 barley str1
100	16	61.5	157	2	Q8LI53 ORYSA	Q8LI53 oryza sativ
101	16	61.5	158	2	Q8CCCG MOUSE	Q8CCCG mus musculu
102	16	61.5	161	2	Q23659 CAEBL	Q23659 caenorhabdi
103	16	61.5	162	2	Q91G16 ADEP3	Q91G16 porcine ade
104	16	61.5	164	2	Q7LZW7 CVB	Q7LZW7 chrysaenthem

105	16	61.5	166	2	061045_DROME	061445_drosophila	178	16	61.5	272	2	0655G8_9CNIID	0655g8_nematostell
106	16	61.5	167	2	024969_GIALA	024969_giardia lam	179	16	61.5	272	2	06WJ39_9CNIID	06wj39_nematostell
107	16	61.5	170	2	09B117_GIALA	09b117_giardia lam	180	16	61.5	276	2	04W123_ASPFU	04w123_aspergillus
108	16	61.5	170	2	08XS27_RALSTONIA	08xs27_ralstonia s	181	16	61.5	280	2	06ZRL6_HUMAN	06zrl6_homo sapien
109	16	61.5	171	2	06ZP60_HUMAN	06zp60_homo sapien	182	16	61.5	280	2	06ZUB8_HUMAN	06zub8_homo sapien
110	16	61.5	171	2	060ZL5_CABER	060zl5_caenorhabdi	183	16	61.5	281	2	07ZHQ7_9H1V1	07zhq7_human immun
111	16	61.5	171	2	08MOA0_CABEL	08moa0_caenorhabdi	184	16	61.5	282	1	KR101_HUMAN	P60331_homo sapien
112	16	61.5	171	2	05S8S1_PSEFU	05s8s1_pseudomonas	185	16	61.5	284	2	06MIV8_DROME	06miv8_drosophila
113	16	61.5	172	2	0988Y7_RHILLO	0988y7_rhizobium l	186	16	61.5	285	2	08S902_SOYBN	08s902_glycine max
114	16	61.5	174	2	071D98_DROYA	071d98_drosophila	187	16	61.5	287	2	08WVJ7_PASCT	08wvj7_bottemia vi
115	16	61.5	175	2	09VS07_DROME	09vs07_xanthomonas	188	16	61.5	287	2	04X7N6_PLACH	04x7n6_plasmodium
116	16	61.5	175	2	04JMM7_9BACT	04jmm7_drosophila	189	16	61.5	287	2	05U140_DROME	05u140_drosophila
117	16	61.5	179	2	04M046_9BURK	04m046_burkholderi	190	16	61.5	290	2	06EP14_DROME	06ep14_drosophila
118	16	61.5	181	2	06ZTR4_HUMAN	06ztr4_homo sapien	191	16	61.5	290	2	04EX06_LEITMA	04ex06_leishmania
119	16	61.5	188	2	04RZL3_TETNG	04rzl3_tetradodon n	192	16	61.5	291	2	06RHX7_LYCES	06rhx7_lycopersico
120	16	61.5	189	2	05GXC7_XANOR	05gxc7_xanthomonas	193	16	61.5	291	2	05B3N4_EBENT	05b3n4_aspergillus
121	16	61.5	191	2	0700R6_GIALA	0700r6_giardia lam	194	16	61.5	294	2	06Z0N2_CABER	06z0n2_caenorhabdi
122	16	61.5	192	2	080A8_TOSV	080a8_toscana vir	195	16	61.5	294	2	09XUS0_CABEL	09xus0_caenorhabdi
123	16	61.5	203	2	06ZYU4_HUMAN	06zyu4_homo sapien	196	16	61.5	294	2	04RJG2_TETNG	04rjg2_tetradodon n
124	16	61.5	206	2	04SKY8_TETNG	04sky8_tetradodon n	197	16	61.5	296	2	07X9B9_ARATH	07x9b9_arabidopsis
125	16	61.5	213	2	074ZG8_ASHGO	074zg8_ashbya goss	198	16	61.5	298	2	08H5F5_ORYSA	08h5f5_oryza sativ
126	16	61.5	214	2	09W3F3_DROME	09w3f3_drosophila	199	16	61.5	300	2	04SUT2_TETNG	04sut2_tetradodon n
127	16	61.5	215	2	07R428_GIALA	07r428_giardia lam	200	16	61.5	306	2	084566_CHLTR	084566_chlamydia t
128	16	61.5	215	2	060UH4_9GAMM	060uh4_methylomona	201	16	61.5	309	2	04STZ8_TETNG	04stz8_tetradodon n
129	16	61.5	215	2	060UH6_9GAMM	060uh6_methylomona	202	16	61.5	311	2	051PB5_MAGGR	051pb5_magnaporthe
130	16	61.5	216	2	05TV10_ANOGA	05tv10_anopheles g	203	16	61.5	312	1	09ZUS6_ARATH	09zun8_arabidopsis
131	16	61.5	218	2	08G4A2_BIFLO	08g4a2_bifidobacte	204	16	61.5	321	2	06AU03_ORYSA	06au03_oryza sativ
132	16	61.5	219	2	063SB2_BURPS	063sb2_burkholderi	205	16	61.5	335	2	0727E2_DBSVA	0727e2_desulfotobr
133	16	61.5	221	2	05WNK7_CABER	05wnk7_caenorhabdi	206	16	61.5	335	2	07X1L8_ORYSA	07x1l8_oryza sativ
134	16	61.5	222	2	04VY8_DROME	04vy8_drosophila	207	16	61.5	340	2	07ZP51_9H1V1	07zp51_oryza sativ
135	16	61.5	225	2	027014_TENMO	027014_tenebrio mo	208	16	61.5	346	2	07ZP51_9H1V1	07zp51_human immun
136	16	61.5	225	2	027013_TENMO	027013_tenebrio mo	209	16	61.5	347	2	04TW08_9H1V1	04tw08_human immun
137	16	61.5	228	2	04VSS4_DROME	04vss4_drosophila	210	16	61.5	347	2	04QCP8_LEITMA	04qcp8_leishmania
138	16	61.5	232	2	06Z1R3_BURMA	06z1r3_burkholderi	211	16	61.5	349	1	PAK4_PAT	08g38_rattus nov
139	16	61.5	233	2	08LHK2_ORYSA	08lhk2_oryza sativ	212	16	61.5	349	2	04TW05_9H1V1	04tw05_human immun
140	16	61.5	237	2	06YU04_ORYSA	06yu04_oryza sativ	213	16	61.5	351	2	04TW09_9H1V1	04tw09_human immun
141	16	61.5	238	2	070U60_GIALA	070u60_giardia lam	214	16	61.5	353	2	067VP8_ORYSA	067vp8_oryza sativ
142	16	61.5	240	2	093877_FUSOX	093877_fusarium ox	215	16	61.5	353	2	04TWU7_9H1V1	04twu7_human immun
143	16	61.5	241	2	09VVW4_DROME	09vvw4_drosophila	216	16	61.5	354	2	04TWV0_9H1V1	04twv0_human immun
144	16	61.5	242	2	004701_NECBA	004701_nectria hae	217	16	61.5	355	2	08M1T0_BOVIN	08m1t0_bos taurus
145	16	61.5	243	2	05QHS3_BRABE	05qhs3_brachydanio	218	16	61.5	356	2	04TWV2_9H1V1	04twv2_human immun
146	16	61.5	245	1	CTRR_GADMO	P80646_gadus morhu	219	16	61.5	358	2	053XJ7_ORYSA	053xj7_oryza sativ
147	16	61.5	247	2	018458_HETGL	018458_heterodera	220	16	61.5	362	2	08CSN9_MOUSE	08csn9_mus musculu
148	16	61.5	249	2	05VRX0_ORYSA	05vrx0_oryza sativ	221	16	61.5	363	2	05H1J8_XANOR	05h1j8_xanthomonas
149	16	61.5	251	2	024774_ENCBU	024774_encyttraeus	222	16	61.5	366	2	05L1J3_SILPO	05l1j3_silicibacte
150	16	61.5	251	2	070LO4_ENCBU	070lo4_encyttraeus	223	16	61.5	370	2	07XEN8_ORYSA	07xen8_oryza sativ
151	16	61.5	252	1	TRY1_DROME	P52905_drosophila	224	16	61.5	371	2	04P4V1_USTMA	04p4v1_usciliago ma
152	16	61.5	252	2	0540Z6_DROME	0540z6_drosophila	225	16	61.5	373	2	06ABK7_PROAC	06abk7_propionibac
153	16	61.5	254	2	060K89_CABER	060k89_caenorhabdi	226	16	61.5	373	2	05SNS5_BRABE	05sns5_birachydanio
154	16	61.5	258	2	09YOB3_DROME	09yob3_drosophila	227	16	61.5	374	2	07PHB7_ANOGA	07phb7_caenorhabdi
155	16	61.5	259	2	09GQ40_GIALA	09gq40_giardia lam	228	16	61.5	378	2	0721R1_CABEL	0721r1_caenorhabdi
156	16	61.5	259	2	058PZ0_ARATH	058pzo_arabidopsis	229	16	61.5	378	2	094699_PABPR	094699_parametium
157	16	61.5	259	2	08H1N1_ARATH	08h1n1_arabidopsis	230	16	61.5	376	2	064CRI_9ARCH	064crl_trypanosoma
158	16	61.5	259	2	06ACZ2_BRABE	06acz2_brachydanio	231	16	61.5	377	2	057WD2_9TRYP	057wd2_cryptosporidia
159	16	61.5	260	2	06ZPV9_XENTR	06zpv9_xenopus tro	232	16	61.5	378	2	063RGO_BURPS	063rgo_burkholderi
160	16	61.5	260	2	09W7Q3_PAROL	09w7q3_paralichthy	233	16	61.5	380	2	062166_BURMA	062166_burkholderi
161	16	61.5	261	2	06DHD9_BRABE	06dhd9_canine ente	234	16	61.5	380	2	062166_BURMA	062166_burkholderi
162	16	61.5	262	1	072113_9H1V1	072113_human immun	235	16	61.5	382	2	0970G3_ARATH	0970g3_arabidopsis
163	16	61.5	262	1	VME1_CVCAI	P36299_canine ente	236	16	61.5	382	2	08XW74_CLOPE	08xw74_clostridium
164	16	61.5	262	2	053KW3_ORYSA	053kw3_oryza sativ	237	16	61.5	382	2	04W993_ASPFU	04w993_aspergillus
165	16	61.5	263	1	CTRA_GADMO	P47799_gadus morhu	238	16	61.5	386	2	04QB26_LEITMA	04qb26_leishmania
166	16	61.5	263	2	09PWQ6_GADMO	09pwq6_gadus morhu	239	16	61.5	388	2	06Z7H7_CABER	06z7h7_caenorhabdi
167	16	61.5	263	2	04SEW1_TETNG	04sew1_tetradodon n	240	16	61.5	392	2	06P6D1_HUMAN	06p6d1_homo sapien
168	16	61.5	263	2	05H2D0_XENTR	05h2d0_xenopus tro	241	16	61.5	394	2	09GQ47_GIALA	09gq47_giardia lam
169	16	61.5	263	2	051029_XENTR	051029_xenopus tiro	242	16	61.5	397	2	0623N1_CABER	0623n1_caenorhabdi
170	16	61.5	263	2	06GNF7_XENLA	06gnf7_xenopus lae	243	16	61.5	397	2	09XUB6_CABEL	09xub6_caenorhabdi
171	16	61.5	263	2	06PGS4_XENLA	06pgs4_xenopus lae	244	16	61.5	399	2	07SET6_NEUCRA	07set6_neurospora
172	16	61.5	263	2	07BSY4_XENLA	07bsy4_xenopus lae	245	16	61.5	400	2	063R93_BURPS	063r93_burkholderi
173	16	61.5	264	2	086EC4_SCHUA	086ec4_schistosoma	246	16	61.5	401	2	07PV94_ANOGA	07pv94_anopheles g
174	16	61.5	264	2	04QY78_SPAUV	04qy78_sparus aura	247	16	61.5	404	2	09C211_NEUCR	09c211_neutrospora
175	16	61.5	265	2	05GIU4_BRABE	05giu4_brachydanio	248	16	61.5	404	2	04RPNO_TETNG	04rpn0_tetradodon n
176	16	61.5	266	2	05MS69_LEGPA	05ms69_legionella	249	16	61.5	410	2	05H2U4_XENTR	05h2u4_xenopus tro
177	16	61.5	271	2	Q4S096_TETNG	Q4s096_tetradodon n	250	16	61.5	411	2	095SV1_DROME	095sv1_drosophila

251	16	61.5	411	2	Q9VWV5_DROME	Q9VWV5_drosophila	324	16	61.5	572	2	Q42728_RHO70	Q42728_rhodospirid
252	16	61.5	415	2	OSKC03_CRVNE	OSKC03_cryptococcus	325	16	61.5	574	1	GAGJ_DROFY	P21331_drosophila
253	16	61.5	415	2	OSTVP2_ANOGA	OSTVP2_anopheles g	326	16	61.5	574	1	OSX7Q4_BOTCI	OSX7Q4_botrytis ci
254	16	61.5	415	2	Q7V960_FROHM	Q7V960_prochloroc	327	16	61.5	574	2	Q4RBL1_MACPA	Q4RBL1_maccia fasc
255	16	61.5	417	2	Q4Q802_LEIMA	Q4Q802_leishmania	328	16	61.5	581	2	Q96WM9_BOTCI	Q96WM9_botrytis ci
256	16	61.5	420	2	Q4QD22_LEIMA	Q4QD22_leishmania	329	16	61.5	587	2	Q7R1Z0_GIALA	Q7R1Z0_giardia lam
257	16	61.5	421	2	Q6W340_9H1V1	Q6W340_human immun	330	16	61.5	589	2	Q6GQU6_MOUSE	Q6GQU6_mus muscucu
258	16	61.5	422	1	RAD52_CHICK	P39022_gallus gall	331	16	61.5	590	1	IRAK2_HUMAN	Q43187_homo sapien
259	16	61.5	426	2	Q72W52_LEPIC	Q72W52_leptospiro	332	16	61.5	600	1	EGFL5_HUMAN	Q9H144_homo sapien
260	16	61.5	426	2	Q8P9T9_LEPIC	Q8P9T9_leptospiro	333	16	61.5	603	2	OS5K58_9TRYR	OS5K58_hypanosoma
261	16	61.5	430	2	Q840C2_9GAMM	Q840C2_cellulibrio	334	16	61.5	603	2	Q4DHG3_9LAMI	Q4DHG3_pertilla cit
262	16	61.5	436	2	Q26859_TRYCO	Q26859_trypanosoma	335	16	61.5	607	2	OSU388_BRABE	OSU388_brachydanio
263	16	61.5	440	2	Q941N4_AYESA	Q941N4_avena sativ	336	16	61.5	608	2	OS7UW7_9TRYR	OS7UW7_trypanosoma
264	16	61.5	441	2	Q6UAP5_TETNG	Q6UAP5_tetradodon n	337	16	61.5	613	2	Q95TQ2_DROME	Q95TQ2_drosophila
265	16	61.5	443	2	Q6AZB6_BRARE	Q6AZB6_brachydanio	338	16	61.5	613	2	Q9V5F6_DROME	Q9V5F6_drosophila
266	16	61.5	446	2	Q7XGP0_ORISA	Q7XGP0_oryza sativ	339	16	61.5	616	2	Q738S5_BACCI	Q738S5_bacillus ce
267	16	61.5	446	2	Q8SB64_ORISA	Q8SB64_oryza sativ	340	16	61.5	622	2	Q7QMD2_GIALA	Q7QMD2_giardia lam
268	16	61.5	447	2	Q8TL45_METAC	Q8TL45_methanosarc	341	16	61.5	625	2	Q5K546_HUMAN	Q5K546_homo sapien
269	16	61.5	448	2	Q5GVX5_XANOR	Q5GVX5_xanthomonas	342	16	61.5	625	2	Q5R810_PONPY	Q5R810_pongo pygma
270	16	61.5	450	2	Q7XP30_ORISA	Q7XP30_oryza sativ	343	16	61.5	629	2	Q9DWH3_RCMVM	Q9DWH3_rat cytoomeg
271	16	61.5	452	2	Q8O000_METMA	Q8O000_methanosarc	344	16	61.5	632	2	Q6YXW2_ORISA	Q6YXW2_oryza sativ
272	16	61.5	453	2	Q20490_CABEL	Q20490_caenorhabdi	345	16	61.5	633	2	Q8WTK1_9H1V1	Q8WTK1_meloidogyne
273	16	61.5	455	2	Q4PT10_ARATH	Q4PT10_arabidopsis	346	16	61.5	636	2	Q61BF2_CABBR	Q61BF2_caenorhabdi
274	16	61.5	455	2	Q6IFV6_RAT	Q6IFV6_rattus norv	347	16	61.5	638	1	YT32_CABEL	Q20680_caenorhabdi
275	16	61.5	457	2	Q9YQZ1_9HERP	Q9YQZ1_randia herpe	348	16	61.5	639	2	Q84K89_AYESA	Q84K89_avena sativ
276	16	61.5	461	2	Q4Q6C9_LEIMA	Q4Q6C9_leishmania	349	16	61.5	642	2	Q9PPI1_CAMJE	Q9PPI1_campylobact
277	16	61.5	463	2	Q6Z3D6_ORISA	Q6Z3D6_oryza sativ	350	16	61.5	645	2	Q97448_GIALA	Q97448_giardia lam
278	16	61.5	464	2	Q5CTB1_CRYPV	Q5CTB1_cryptospori	351	16	61.5	652	2	Q4S9Q3_TETNG	Q4S9Q3_tetradodon n
279	16	61.5	464	2	Q5CIV7_CRYHO	Q5CIV7_cryptospori	352	16	61.5	660	2	Q7QY47_GIALA	Q7QY47_giardia lam
280	16	61.5	466	2	Q5U8X1_9H1V1	Q5U8X1_human immun	353	16	61.5	660	2	Q6PCHO_XENILA	Q6PCHO_xenopus lae
281	16	61.5	466	2	Q8GY60_ARATH	Q8GY60_arabidopsis	354	16	61.5	665	1	YDER_SCHPO	Q10447_schizosacch
282	16	61.5	470	2	Q8ZEG0_STRAM	Q8ZEG0_streptomyce	355	16	61.5	679	2	Q9XTR3_GIALA	Q9XTR3_giardia lam
283	16	61.5	471	2	Q5TYV3_MOCRA	Q5TYV3_nocardia fa	356	16	61.5	681	2	Q6LBN7_BOVIN	Q6LBN7_bos taurus
284	16	61.5	473	2	Q5ZCJ2_MAGGR	Q5ZCJ2_magnaporthe	357	16	61.5	687	2	Q817G1_CABEL	Q817G1_caenorhabdi
285	16	61.5	473	2	Q5FAD8_NEIG1	Q5FAD8_neisseria g	358	16	61.5	695	1	TRFL_HORSE	Q78111_equus caball
286	16	61.5	473	2	Q9JWV9_NEIMA	Q9JWV9_neisseria m	359	16	61.5	696	2	Q5R2V7_HUMAN	Q5R2V7_homo sapien
287	16	61.5	473	2	Q9JXZ7_NEIMB	Q9JXZ7_neisseria m	360	16	61.5	699	2	Q93HN9_STRAM	Q93HN9_streptomyce
288	16	61.5	475	2	Q9X8J5_STRCO	Q9X8J5_streptomyce	361	16	61.5	708	1	TRFL_BOVIN	P24627_bos taurus
289	16	61.5	479	2	Q4SLI6_TETNG	Q4SLI6_tetradodon n	362	16	61.5	708	1	TRFL_BUBBU	Q77698_bubalus bub
290	16	61.5	480	2	Q7SZR0_BRARE	Q7SZR0_brachydanio	363	16	61.5	708	1	TRFL_CAMDR	Q97un0_camelus dro
291	16	61.5	481	2	Q9D635_MOUSE	Q9D635_mus muscucu	364	16	61.5	708	1	TRFL_CAPHI	Q29477_capra hircu
292	16	61.5	486	2	Q9LQ21_ARATH	Q9LQ21_arabidopsis	365	16	61.5	708	1	Q6LEC7_BOVIN	Q6LEC7_bos taurus
293	16	61.5	488	2	Q5WMW6_LEGPL	Q5WMW6_legionella	366	16	61.5	711	1	TRFL_HUMAN	P02788_homo sapien
294	16	61.5	488	2	Q5AL61_CANAL	Q5AL61_candida alb	367	16	61.5	711	2	Q5D5M0_HUMAN	Q5D5M0_homo sapien
295	16	61.5	489	2	Q4HYK7_9DELT	Q4HYK7_aeromonas	368	16	61.5	711	2	Q6FBS7_ACTAD	Q6FBS7_actinobact
296	16	61.5	501	2	Q4HVA7_GIBZE	Q4HVA7_gibberella	369	16	61.5	713	2	Q5EKS1_HUMAN	Q5EKS1_homo sapien
297	16	61.5	510	2	MS2P_CRIGR	Q54862_cricetulus	370	16	61.5	719	1	ERGT_PNECA	Q75775_ashyba gos
298	16	61.5	513	1	Q4IVN9_AZOV1	Q4IVN9_azotobacter	371	16	61.5	719	2	Q757T5_ASHGO	Q757T5_pneumocysti
299	16	61.5	517	2	Q4VBN3_RAT	Q4VBN3_rattus norv	372	16	61.5	722	2	Q6F822_CANCA	Q6F822_candida gla
300	16	61.5	520	2	Q8NCV0_HUMAN	Q8NCV0_homo sapien	373	16	61.5	734	2	Q9NSE6_HUMAN	Q9NSE6_homo sapien
301	16	61.5	520	2	Q8IYS4_HUMAN	Q8IYS4_homo sapien	374	16	61.5	740	2	Q5VYP8_9TRYR	Q5VYP8_trypanosoma
302	16	61.5	522	2	Q53P11_ORISA	Q53P11_oryza sativ	375	16	61.5	741	2	Q4QD80_LEIMA	Q4QD80_leishmania
303	16	61.5	523	2	Q86MK1_DROME	Q86MK1_drosophila	376	16	61.5	743	2	Q4QD78_LEIMA	Q4QD78_leishmania
304	16	61.5	528	2	Q53NH3_ORISA	Q53NH3_oryza sativ	377	16	61.5	777	2	Q4QD79_LEIMA	Q4QD79_leishmania
305	16	61.5	532	2	Q8XIC0_USYMA	Q8XIC0_ustilago ma	378	16	61.5	777	2	Q9FEV4_STRCO	Q9FEV4_streptomyce
306	16	61.5	532	2	Q4PEV3_USYMA	Q4PEV3_ustilago ma	379	16	61.5	783	2	Q76L28_9VIRU	Q76L28_helicobasid
307	16	61.5	533	2	Q74EC9_GEOSL	Q74EC9_geobacter s	380	16	61.5	788	2	Q5NM24_9ARCH	Q5NM24_uncultured
308	16	61.5	536	2	Q4U3U6_NEUAF	Q4U3U6_neutrospora	381	16	61.5	795	2	Q4IEN4_GIBZE	Q4IEN4_gibberella
309	16	61.5	536	2	Q4T8X5_TETNG	Q4T8X5_tetradodon n	382	16	61.5	802	2	Q4ID36_GIBZE	Q4ID36_gibberella
310	16	61.5	537	1	QAY_NEUCR	P16166_neutrospora	383	16	61.5	810	2	Q5VQ42_ORISA	Q5VQ42_oryza sativ
311	16	61.5	537	1	Q5BAQ4_EMENT	Q5BAQ4_aspergillus	384	16	61.5	825	2	Q4Q608_LEIMA	Q4Q608_leishmania
312	16	61.5	538	2	Q6ZIB7_ORISA	Q6ZIB7_oryza sativ	385	16	61.5	825	2	Q7RPF7_PLAYO	Q7RPF7_plasmodium
313	16	61.5	542	2	Q9FP95_ARATH	Q9FP95_arabidopsis	386	16	61.5	842	2	Q61IW9_CABBR	Q61IW9_caenorhabdi
314	16	61.5	542	2	Q86ZM8_PODAN	Q86ZM8_podospora a	387	16	61.5	843	2	Q8P124_XANAC	Q8P124_xanthomonas
315	16	61.5	545	2	Q4IAQ1_GIBZE	Q4IAQ1_gibberella	388	16	61.5	845	2	Q8AK00_9H1V1	Q8AK00_human immun
316	16	61.5	546	2	Q5M281_9AGAR	Q5M281_pleurocous s	389	16	61.5	845	2	Q4Q607_LEIMA	Q4Q607_leishmania
317	16	61.5	549	2	Q5IXV6_MAGGR	Q5IXV6_magnaporthe	390	16	61.5	847	2	Q8AP19_9H1V1	Q8AP19_human immun
318	16	61.5	552	2	Q4IRQ1_GIBZE	Q4IRQ1_gibberella	391	16	61.5	847	2	Q4QWT9_9H1V1	Q4QWT9_human immun
319	16	61.5	552	2	Q6IDJ9_XENTR	Q6IDJ9_xenopus cto	392	16	61.5	848	2	Q4QW55_9H1V1	Q4QW55_human immun
320	16	61.5	560	2	Q5B514_EMENT	Q5B514_aspergillus	393	16	61.5	851	2	Q7RVV6_NEUCR	Q7RVV6_neutrospora
321	16	61.5	561	2	Q96WN0_BOTCI	Q96WN0_botrytis ci	394	16	61.5	852	2		
322	16	61.5					395	16	61.5				
323	16	61.5					396	16	61.5				

397	16	61.5	852	2	Q9HB44_NEUCR	Q9h44	neuropora	470	16	61.5	2128	2	Q5ONG8_ENTHT	Q5ong8	entamoeba h
398	16	61.5	852	2	Q70010_9H1V1	Q70010	human immun	471	16	61.5	2233	2	Q94711_PARTE	Q94711	paramecium
399	16	61.5	854	2	Q5BEC9_9H1V1	Q5ec9	human immun	472	16	61.5	2273	2	Q5ICB9_macCPA	Q5icb9	macaca faec
400	16	61.5	858	2	Q8AB09_9H1V1	Q8ab09	human immun	473	16	61.5	2377	2	Q7WQ80_BORBR	Q7wq80	bordelella
401	16	61.5	859	2	Q4Q027_LEIMA	Q4q027	leishmania	474	16	61.5	2395	2	Q27167_PARTE	Q27167	paramecium
402	16	61.5	859	2	Q8ADX7_9H1V1	Q8adx7	human immun	475	16	61.5	2397	2	Q6Q7Y4_PARPR	Q6q7y4	paramecium
403	16	61.5	859	2	Q8UNP8_9H1V1	Q8unp8	human immun	476	16	61.5	2533	2	P90589_PARTE	P90589	paramecium
404	16	61.5	860	2	Q998E5_9H1V1	Q998e5	human immun	477	16	61.5	2533	2	Q27183_PARTE	Q27183	paramecium
405	16	61.5	864	2	Q4U536_9H1V1	Q4u536	human immun	478	16	61.5	2543	2	P90649_PARPR	P90649	paramecium
406	16	61.5	865	2	Q73342_9H1V1	Q73342	human immun	479	16	61.5	2715	1	G156_PARPR	G156	paramecium
407	16	61.5	865	2	Q73353_9H1V1	Q73353	human immun	480	16	61.5	2986	1	Q4HFM4_GIBBZ	Q4hfm4	glibbze
408	16	61.5	870	2	Q5EBD7_9H1V1	Q5ebd7	human immun	481	16	61.5	3596	1	RPOA_SHPV	RPOA	shpv
409	16	61.5	877	2	Q5VG71_9H1V1	Q5vg71	human immun	482	16	61.5	3869	2	Q5CRCO_CRYPV	Q5crco	cryptospori
410	16	61.5	879	2	Q4M9A0_ASFPV	Q4m9a0	aspergillus	483	16	61.5	3869	2	Q86PQ3_CRYPV	Q86pq3	cryptospori
411	16	61.5	879	2	Q4PQC4_USITMA	Q4pqc4	usitlago ma	484	16	61.5	4382	2	Q5DIC6_CVHSA	Q5dic6	sars corona
412	16	61.5	881	2	Q4N476_THBPA	Q4n476	theliera p	485	16	61.5	4382	2	Q5Y188_CVHSA	Q5y188	sars corona
413	16	61.5	884	2	Q80718_9H1V1	Q8q718	human immun	486	16	61.5	4382	2	Q692B5_CVHSA	Q692b5	sars corona
414	16	61.5	889	2	Q8IUQ3_PLAF7	Q8i1q3	plasmodium	487	16	61.5	4382	2	Q6JH33_CVHSA	Q6jh39	sars corona
415	16	61.5	897	2	Q5SSQ4_CRYNE	Q5s04	cryptococcu	488	16	61.5	4382	2	Q6JH47_CVHSA	Q6jh47	sars corona
416	16	61.5	897	2	Q5KGF6_CRYNE	Q5k5f6	cryptococcu	489	16	61.5	4382	2	Q6R7Y7_CVHSA	Q6r7y7	sars corona
417	16	61.5	907	2	Q5FM49_ARATH	Q5fm49	arabidopsis	490	16	61.5	4382	2	Q6RCM6_CVHSA	Q6rcm6	sars corona
418	16	61.5	911	2	Q80TU8_MOUSE	Q80tj8	mus musculu	491	16	61.5	4382	2	Q6RCX7_CVHSA	Q6rcx7	sars corona
419	16	61.5	912	2	Q4JHL4_FUGRU	Q4jhl4	fugu rubrip	492	16	61.5	4382	2	Q6RCY8_CVHSA	Q6rcy8	sars corona
420	16	61.5	918	2	Q6ETC9_9URUC	Q6etj9	icopileura	493	16	61.5	4382	2	Q6RCZ9_CVHSA	Q6rcz9	sars corona
421	16	61.5	920	2	Q7XGN4_ORYSA	Q7xgn4	oryza sativ	494	16	61.5	4382	2	Q6RD10_CVHSA	Q6rd10	sars corona
422	16	61.5	920	2	Q94GU3_ORYSA	Q94gu3	oryza sativ	495	16	61.5	4382	2	Q6RD21_CVHSA	Q6rd21	sars corona
423	16	61.5	929	2	Q5EBD8_9H1V1	Q5ebd8	human immun	496	16	61.5	4382	2	Q6RD32_CVHSA	Q6rd32	sars corona
424	16	61.5	934	2	Q87R90_VIBPA	Q87r90	vibrio para	497	16	61.5	4382	2	Q6RD43_CVHSA	Q6rd43	sars corona
425	16	61.5	935	2	Q4P9H9_USITMA	Q4p9h9	usitlago ma	498	16	61.5	4382	2	Q6RD54_CVHSA	Q6rd54	sars corona
426	16	61.5	939	2	Q5RHW8_BRABE	Q5rhw8	brachyolma	499	16	61.5	4382	2	Q6RD65_CVHSA	Q6rd65	sars corona
427	16	61.5	967	2	Q4T490_TETNG	Q4t490	tetradion n	500	16	61.5	4382	2	Q6VA79_CVHSA	Q6va79	sars corona
428	16	61.5	975	2	P91357_CABEL	P91357	caenorhabdi	501	16	61.5	4382	2	Q6VA90_CVHSA	Q6va90	sars corona
429	16	61.5	999	1	SMG_DROME	Q23372	drosophila	502	16	61.5	4382	2	Q6VAA1_CVHSA	Q6vaa1	sars corona
430	16	61.5	1031	2	Q65500_ARATH	Q65500	arabidopsis	503	16	61.5	4382	2	Q6WGN0_CVHSA	Q6wgn0	sars corona
431	16	61.5	1038	2	Q583N6_9TRYP	Q583n6	trypanosoma	504	16	61.5	5560	1	SPEN_DROME	SPEN	drome
432	16	61.5	1039	2	Q4P295_USITMA	Q4p295	usitlago ma	505	16	61.5	6880	2	Q6S8D8_CVHSA	Q6s8d8	sars corona
433	16	61.5	1069	2	Q592X1_CANAL	Q592x1	candida alb	506	16	61.5	7073	1	RIAB_CVHSA	RIAB	cvhsa
434	16	61.5	1069	2	Q9UJF2_CABEL	Q9ujf2	caenorhabdi	507	16	61.5	7073	2	Q5DIC7_CVHSA	Q5dic7	sars corona
435	16	61.5	1071	2	Q61ZG1_CABEL	Q61zgi	caenorhabdi	508	16	61.5	7073	2	Q5Y189_CVHSA	Q5y189	sars corona
436	16	61.5	1105	2	Q20371_CABEL	Q20371	caenorhabdi	509	16	61.5	7073	2	Q6Y2B6_CVHSA	Q6y2b6	sars corona
437	16	61.5	1136	2	Q92C14_LISIN	Q92c14	lietexia in	510	16	61.5	7073	2	Q6JH40_CVHSA	Q6jh40	sars corona
438	16	61.5	1154	2	Q9GQ46_GIALA	Q9gq46	giardia lam	511	16	61.5	7073	2	Q6JH48_CVHSA	Q6jh48	sars corona
439	16	61.5	1168	2	Q6AVU1_ORYSA	Q6av11	oryza sativ	512	16	61.5	7073	2	Q6R7Y8_CVHSA	Q6r7y8	sars corona
440	16	61.5	1191	2	Q8TUQ3_DROME	Q8tuq3	drosophila	513	16	61.5	7073	2	Q6RPB9_CVHSA	Q6rpb9	sars corona
441	16	61.5	1194	2	Q4FX33_LEIMA	Q4fx33	leishmania	514	16	61.5	7073	2	Q6UZP1_CVHSA	Q6uzp1	sars corona
442	16	61.5	1194	2	Q7FM68_RAT	Q7fm68	rattus norv	515	16	61.5	7073	2	Q6UZP5_CVHSA	Q6uzp5	sars corona
443	16	61.5	1227	2	Q6IFU2_SCHMA	Q6ifu2	schistosoma	516	16	61.5	7073	2	Q6V5B6_CVHSA	Q6v5b6	sars corona
444	16	61.5	1243	2	Q8QTA3_WSSV	Q8qta3	white spot	517	16	61.5	7073	2	Q6RCW7_CVHSA	Q6rcw7	sars corona
445	16	61.5	1243	2	Q8VAF5_WSSV	Q8vaf5	white spot	518	16	61.5	7073	2	Q6RCX8_CVHSA	Q6rcx8	sars corona
446	16	61.5	1243	2	Q91LM1_WSSV	Q91lm1	white spot	519	16	61.5	7073	2	Q6RCY9_CVHSA	Q6rcy9	sars corona
447	16	61.5	1270	2	Q5UXR6_HUMAN	Q5uxr6	homo sapien	520	16	61.5	7073	2	Q6RD00_CVHSA	Q6rd00	sars corona
448	16	61.5	1310	1	GPI25_MOUSE	Q7tc36	mus musculu	521	16	61.5	7073	2	Q6RD11_CVHSA	Q6rd11	sars corona
449	16	61.5	1318	2	Q4QDAS_LEIMA	Q4qdas	leishmania	522	16	61.5	7073	2	Q6RD22_CVHSA	Q6rd22	sars corona
450	16	61.5	1332	2	Q4FPD5_LEIMA	Q4fpd5	leishmania	523	16	61.5	7073	2	Q6RD33_CVHSA	Q6rd33	sars corona
451	16	61.5	1353	2	Q4QBR1_LEIMA	Q4qbr1	leishmania	524	16	61.5	7073	2	Q6RD44_CVHSA	Q6rd44	sars corona
452	16	61.5	1363	2	Q4LIW7_9BURK	Q4liw7	burkholderi	525	16	61.5	7073	2	Q6RD55_CVHSA	Q6rd55	sars corona
453	16	61.5	1367	2	Q4QRT6_LEIMA	Q4qrt6	leishmania	526	16	61.5	7073	2	Q6RD66_CVHSA	Q6rd66	sars corona
454	16	61.5	1369	2	Q4QAT7_LEIMA	Q4qat7	leishmania	527	16	61.5	7073	2	Q6VA80_CVHSA	Q6va80	sars corona
455	16	61.5	1380	1	CYAA_LEIDO	Q27675	leishmania	528	16	61.5	7073	2	Q6VA91_CVHSA	Q6va91	sars corona
456	16	61.5	1467	2	Q61ON6_CABEL	Q61on6	caenorhabdi	529	16	61.5	7073	2	Q6VAA2_CVHSA	Q6vaa2	sars corona
457	16	61.5	1474	2	Q625Q4_CABEL	Q625q4	caenorhabdi	530	16	61.5	7073	2	Q4UDQ5_CVHSA	Q4udq5	sars corona
458	16	61.5	1486	2	Q4GDM0_TETNG	Q4gdm0	tetradion n	531	16	61.5	7073	2	Q4UDR8_CVHSA	Q4udr8	sars corona
459	16	61.5	1557	2	Q4QHE6_LEIMA	Q4qhe6	leishmania	532	15	57.7	26	2	Q93940_PODAN	Q93940	podospora a
460	16	61.5	1568	2	Q55TP3_CRYNE	Q55tp3	cryptococcu	533	15	57.7	33	2	Q9BY90_HUMAN	Q9by90	homo sapien
461	16	61.5	1585	2	Q5KIP2_CRYNE	Q5kip2	cryptococcu	534	15	57.7	40	1	MT1_DROME	MT1	drome
462	16	61.5	1755	2	Q4QPD0_LEIMA	Q4qpd0	leishmania	535	15	57.7	40	1	MT1_DROME	MT1	drome
463	16	61.5	1769	2	Q7WC76_BORPA	Q7wc76	bordelella	536	15	57.7	40	1	COAT2_BROPI	COAT2	bropi
464	16	61.5	1818	1	BNC_DROME	Q8mxi	drosophila	537	15	57.7	43	1	COAT2_XANCP	COAT2	xanncp
465	16	61.5	1823	1	Q5BTH9_DROME	Q5bth9	drosophila	538	15	57.7	43	1	Q4UUV8_XANCP	Q4uuv8	xanncp
466	16	61.5	2071	2	Q7KW61_DROME	Q7kw61	drosophila	539	15	57.7	43	2	Q4UUV8_XANCP	Q4uuv8	xanncp
467	16	61.5	2071	2	Q7KW61_DROME	Q7kw61	giardia lam	540	15	57.7	43	2	Q4UUV8_XANCP	Q4uuv8	xanncp
468	16	61.5	2087	2	Q4S488_TETNG	Q4s488	tetradion n	541	15	57.7	46	2	Q8BPK1_XANCP	Q8bpk1	xanncp
469	16	61.5	2117	2	Q4QIS6_LEIMA	Q4qis6	leishmania	542	15	57.7	46	2	Q8BPK1_XANCP	Q8bpk1	xanncp

543	15	57.7	53	2	069338_9GAMA	069338	salimixiline
544	15	57.7	54	2	074959_GEOSL	074959	geobacter s
545	15	57.7	56	2	05C2S7_SCHJA	05C2S7	schistosoma
546	15	57.7	57	2	060CB8_MERCA	060CB8	methylcocc
547	15	57.7	58	2	096814_9BETA	096814	human herpe
548	15	57.7	58	2	096915_9BETA	096915	human herpe
549	15	57.7	58	2	04TUC8_TERNG	04TUC8	tetradon n
550	15	57.7	59	2	07S2A3_NEUCR	07S2A3	neutropora
551	15	57.7	59	2	05CIN4_SCHJA	05CIN4	schistosoma
552	15	57.7	59	2	04PP35_TERNG	04PP35	tetradon n
553	15	57.7	61	2	086TX3_HUMAN	086TX3	homo sapien
554	15	57.7	63	2	06FUN6_CANGA	06FUN6	candida gla
555	15	57.7	64	2	06KOM3_9VIRU	06KOM3	human astro
556	15	57.7	64	2	08JMX7_9VIRU	08JMX7	human astro
557	15	57.7	67	2	04TGES_TERNG	04TGES	tetradon n
558	15	57.7	68	2	04THB0_TERNG	04THB0	tetradon n
559	15	57.7	72	2	05DEB6_ANGCA	05DEB6	anopheles g
560	15	57.7	73	2	07R2J5_GIALA	07R2J5	giardia lam
561	15	57.7	73	2	05RL97_PIG	05RL97	sus scrofa
562	15	57.7	74	2	06L127_BURMA	06L127	burholderi
563	15	57.7	74	2	08YSS5_ANNAP	08YSS5	anabaena ap
564	15	57.7	75	2	05YR21_9VIRU	05YR21	rock bream
565	15	57.7	77	2	04UTZ9_XANCP	04UTZ9	xanthomonas
566	15	57.7	77	2	08PDJ8_XANCP	08PDJ8	xanthomonas
567	15	57.7	78	2	006467_XANCA	006467	xanthomonas
568	15	57.7	79	2	026086_9TIRB	026086	polycelis n
569	15	57.7	79	2	092NT4_RHIME	092NT4	rhicobium m
570	15	57.7	80	2	04J5G5_AZOV1	04J5G5	azotobacter
571	15	57.7	80	2	05NCE5_MOUSE	05NCE5	mus musculus
572	15	57.7	80	2	069752_9H1V1	069752	human immun
573	15	57.7	81	2	080226_9H1V1	080226	human immun
574	15	57.7	82	1	LCR45_ARATH	LCR45	arabidopsis
575	15	57.7	85	2	09BYT5_HUMAN	09BYT5	homo sapien
576	15	57.7	85	2	084GQ2_9SYNE	084GQ2	synecococc
577	15	57.7	85	2	P86K25_9H1V1	P86K25	human immun
578	15	57.7	86	2	06TRY3_CULOU	06TRY3	culex quinq
579	15	57.7	88	2	094203_CLAFOU	094203	cladosporiu
580	15	57.7	88	2	090DH6_9H1V1	090DH6	human immun
581	15	57.7	89	2	043185_HUMAN	043185	homo sapien
582	15	57.7	89	2	051BM5_ENTHI	051BM5	entomoba h
583	15	57.7	89	2	04TDR5_AZOV1	04TDR5	azotobacter
584	15	57.7	90	2	06TVM4_9POXY	06TVM4	orf virus
585	15	57.7	91	2	04XC49_PLACH	04XC49	plasmidom
586	15	57.7	91	2	096Y16_ARATH	096Y16	arabidopsis
587	15	57.7	91	2	039862_9VIRU	039862	human astro
588	15	57.7	91	2	06TMO4_9POXY	06TMO4	orf virus
589	15	57.7	92	2	09UIG9_HUMAN	09UIG9	homo sapien
590	15	57.7	92	2	08SAX8_ORYSA	08SAX8	oryza sativ
591	15	57.7	92	2	P88627_9H1V1	P88627	human immun
592	15	57.7	93	2	05FMFO_LACAC	05FMFO	lactobacill
593	15	57.7	93	2	041098_CHYPI	041098	paramectium
594	15	57.7	93	2	05FZJ8_9HEPC	05FZJ8	hepatitis c
595	15	57.7	93	2	05FZK1_9HEPC	05FZK1	hepatitis c
596	15	57.7	93	2	05FZK8_9HEPC	05FZK8	hepatitis c
597	15	57.7	93	2	078021_9H1V1	078021	human immun
598	15	57.7	94	2	090DH5_9H1V1	090DH5	human immun
599	15	57.7	94	2	09PXN5_9H1V1	09PXN5	human immun
600	15	57.7	95	2	08CSR2_MOUSE	08CSR2	mus musculus
601	15	57.7	96	2	09BYU5_HUMAN	09BYU5	homo sapien
602	15	57.7	96	2	07YTR6_CABEL	07YTR6	caenorhabdi
603	15	57.7	96	2	08VD96_MESAU	08VD96	mesocricetu
604	15	57.7	96	2	077367_9H1V1	077367	human immun
605	15	57.7	96	2	090DH9_9H1V1	090DH9	human immun
606	15	57.7	97	2	08TFH4_MERAC	08TFH4	methanocarc
607	15	57.7	97	2	04Y107_PLABE	04Y107	plasmidom
608	15	57.7	98	2	04UTZ6_XANCP	04UTZ6	xanthomonas
609	15	57.7	98	2	08PDJ3_XANCP	08PDJ3	xanthomonas
610	15	57.7	98	2	090D13_9H1V1	090D13	human immun
611	15	57.7	99	2	05YAU4_AGBOR	05YAU4	agelena ori
612	15	57.7	99	2	071267_9H1V1	071267	human immun
613	15	57.7	99	2	090D11_9H1V1	090D11	human immun
614	15	57.7	99	2	090D12_9H1V1	090D12	human immun
615	15	57.7	100	2	090D15_9H1V1	090D15	human immun
616	15	57.7	101	2	009129_MOUSE	009129	mus musculus
617	15	57.7	101	2	071260_9H1V1	071260	human immun
618	15	57.7	101	2	090DH4_9H1V1	090DH4	human immun
619	15	57.7	101	2	090DH7_9H1V1	090DH7	human immun
620	15	57.7	102	1	VP07_BPAB5	VP07	bacterioph
621	15	57.7	102	1	06MN21_BDEBA	06MN21	bdellovibri
622	15	57.7	102	2	071271_9H1V1	071271	human immun
623	15	57.7	102	2	090D14_9H1V1	090D14	human immun
624	15	57.7	103	2	070RX6_GIALA	070RX6	giardia lam
625	15	57.7	103	2	09UOT5_LOGMT	09UOT5	locusta mig
626	15	57.7	103	2	0806T2_9HEPC	0806T2	hepatitis c
627	15	57.7	104	2	07PG48_ANGCA	07PG48	anopheles g
628	15	57.7	104	2	069W53_ORYSA	069W53	oryza sativ
629	15	57.7	104	2	084BA7_ERWCH	084BA7	erythrina chr
630	15	57.7	106	2	08NOC0_CORGL	08NOC0	corynebacte
631	15	57.7	107	2	07R3Y4_GIALA	07R3Y4	giardia lam
632	15	57.7	107	2	084NA3_9PRBA	084NA3	zapoteca re
633	15	57.7	107	2	09DZX3_9H1V1	09DZX3	human immun
634	15	57.7	107	2	09DZX4_9H1V1	09DZX4	human immun
635	15	57.7	107	2	09DZX5_9H1V1	09DZX5	human immun
636	15	57.7	107	2	09DZX6_9H1V1	09DZX6	human immun
637	15	57.7	107	2	09DZX5_9H1V1	09DZX5	human immun
638	15	57.7	109	1	YGLC_YEAST	YGLC	saccharomy
639	15	57.7	109	2	08CEA7_MOUSE	08CEA7	mus musculus
640	15	57.7	109	2	08JG58_AMBME	08JG58	ambystoma m
641	15	57.7	111	2	06CTJ9_KLUTA	06CTJ9	kluyveromy
642	15	57.7	111	2	092AB2_CHLUP	092AB2	chlamydia p
643	15	57.7	111	2	08C5D5_MOUSE	08C5D5	mus musculus
644	15	57.7	112	2	056714_9HEPC	056714	hepatitis c
645	15	57.7	112	2	056716_9HEPC	056716	hepatitis c
646	15	57.7	112	2	056717_9HEPC	056717	hepatitis c
647	15	57.7	112	2	056718_9HEPC	056718	hepatitis c
648	15	57.7	112	2	056719_9HEPC	056719	hepatitis c
649	15	57.7	112	2	056724_9HEPC	056724	hepatitis c
650	15	57.7	112	2	056726_9HEPC	056726	hepatitis c
651	15	57.7	112	2	056727_9HEPC	056727	hepatitis c
652	15	57.7	112	2	056728_9HEPC	056728	hepatitis c
653	15	57.7	112	2	056729_9HEPC	056729	hepatitis c
654	15	57.7	112	2	056730_9HEPC	056730	hepatitis c
655	15	57.7	112	2	056731_9HEPC	056731	hepatitis c
656	15	57.7	112	2	090DH8_9H1V1	090DH8	human immun
657	15	57.7	113	1	Y011_BPTA	Y011	bacterioph
658	15	57.7	113	2	05AJZ9_CANAL	05AJZ9	candida alb
659	15	57.7	114	2	0614J6_CABRL	0614J6	caenorhabdi
660	15	57.7	114	2	095ZX8_CABEL	095ZX8	caenorhabdi
661	15	57.7	114	2	090D10_9H1V1	090D10	human immun
662	15	57.7	115	2	09S9A1_VISAL	09S9A1	viscum albu
663	15	57.7	116	2	06RUX0_9H1V1	06RUX0	human immun
664	15	57.7	117	2	04NZF5_9DELT	04NZF5	anaeromyxob
665	15	57.7	117	2	08UUD6_9H1V1	08UUD6	human immun
666	15	57.7	117	2	08UUD7_9H1V1	08UUD7	human immun
667	15	57.7	117	2	08UUD8_9H1V1	08UUD8	human immun
668	15	57.7	118	1	CG023_HUMAN	CG023	homo sapien
669	15	57.7	118	2	08RFE3_FUSMN	08RFE3	fruseobacteri
670	15	57.7	119	2	07U993_SYMPX	07U993	synecococc
671	15	57.7	120	2	08UUD9_9H1V1	08UUD9	human immun
672	15	57.7	121	2	069960_9H1V1	069960	human immun
673	15	57.7	121	2	069964_9H1V1	069964	human immun
674	15	57.7	122	1	YK4_CVHSA	YK4	human coron
675	15	57.7	122	2	05KQ12_ORYSA	05KQ12	oryza sativ
676	15	57.7	122	2	05X205_LEGPA	05X205	legionella
677	15	57.7	122	2	05ZS17_LBOPH	05ZS17	leptospira
678	15	57.7	122	2	082209_CHLCTV	082209	chlamydia
679	15	57.7	122	2	05Y181_CVHBA	05Y181	sars corona
680	15	57.7	122	2	0692D8_CVHBA	0692D8	sars corona
681	15	57.7	122	2	06GYO5_CVHBA	06GYO5	sars corona
682	15	57.7	122	2	06R7Y0_CVHBA	06R7Y0	sars corona
683	15	57.7	122	2	06S8D1_CVHBA	06S8D1	sars corona
684	15	57.7	122	2	06S288_CVHSA	06S288	sars corona
685	15	57.7	122	2	06R1D6_CVHSA	06R1D6	sars corona
686	15	57.7	122	2	076XZ9_CVHSA	076XZ9	sars corona
687	15	57.7	122	2	06RCV9_CVHSA	06RCV9	sars corona
688	15	57.7	122	2	06RCX0_CVHSA	06RCX0	sars corona

689	15	57.7	122	2	06RCV1 CVHSA	06rcv1	sars	corona	762	15	57.7	136	2	04V6K1 DROME	04v6k1	drosophila
690	15	57.7	122	2	06RCZ2 CVHSA	06rcz2	sars	corona	763	15	57.7	136	2	06GRP1 ORYSA	06grp1	oryza sativ
691	15	57.7	122	2	06RD03 CVHSA	06rd03	sars	corona	764	15	57.7	137	2	06XUB1 DROPH	06xub1	drosophila
692	15	57.7	122	2	06RD14 CVHSA	06rd14	sars	corona	765	15	57.7	137	2	063U00 BURPS	063u00	burkholderi
693	15	57.7	122	2	06RD25 CVHSA	06rd25	sars	corona	766	15	57.7	137	2	05FXE1 HEPPC	05fxe1	hepatitis c
694	15	57.7	122	2	06RD36 CVHSA	06rd36	sars	corona	767	15	57.7	137	2	05FXE2 HEPPC	05fxe2	hepatitis c
695	15	57.7	122	2	06RD47 CVHSA	06rd47	sars	corona	768	15	57.7	137	2	05FXF2 HEPPC	05fxf2	hepatitis c
696	15	57.7	122	2	06RD58 CVHSA	06rd58	sars	corona	769	15	57.7	137	2	05FXF8 HEPPC	05fxf8	hepatitis c
697	15	57.7	122	2	06SR05 CVHSA	06sr05	sars	corona	770	15	57.7	137	2	05FXG5 HEPPC	05fxg5	hepatitis c
698	15	57.7	122	2	06SR05 CVHSA	06sr05	sars	corona	771	15	57.7	137	2	05FXG9 HEPPC	05fxg9	hepatitis c
699	15	57.7	122	2	06SRR5 CVHSA	06srr5	sars	corona	772	15	57.7	137	2	05FXH0 HEPPC	05fxh0	hepatitis c
700	15	57.7	122	2	06SRX0 CVHSA	06srx0	sars	corona	773	15	57.7	137	2	05FXH1 HEPPC	05fxh1	hepatitis c
701	15	57.7	122	2	06SRS5 CVHSA	06srs5	sars	corona	774	15	57.7	137	2	05FXH2 HEPPC	05fxh2	hepatitis c
702	15	57.7	122	2	06SR70 CVHSA	06sr70	sars	corona	775	15	57.7	137	2	05FXH3 HEPPC	05fxh3	hepatitis c
703	15	57.7	122	2	06SRT5 CVHSA	06srt5	sars	corona	776	15	57.7	137	2	05FXH4 HEPPC	05fxh4	hepatitis c
704	15	57.7	122	2	06SRU0 CVHSA	06sru0	sars	corona	777	15	57.7	137	2	05FXH5 HEPPC	05fxh5	hepatitis c
705	15	57.7	122	2	06SRU5 CVHSA	06sru5	sars	corona	778	15	57.7	137	2	05FXH6 HEPPC	05fxh6	hepatitis c
706	15	57.7	122	2	06SRV0 CVHSA	06srv0	sars	corona	779	15	57.7	137	2	05FXH7 HEPPC	05fxh7	hepatitis c
707	15	57.7	122	2	06SRV5 CVHSA	06srv5	sars	corona	780	15	57.7	137	2	04S169 TETNG	04s169	tetradon n
708	15	57.7	122	2	06SRW0 CVHSA	06srw0	sars	corona	781	15	57.7	138	2	06Y9V8 LYCRS	06y9v8	lycopersico
709	15	57.7	122	2	06SRW5 CVHSA	06srw5	sars	corona	782	15	57.7	138	2	07FEB7 RACMG	07feb7	ribes cyto
710	15	57.7	122	2	06SRX0 CVHSA	06srx0	sars	corona	783	15	57.7	139	2	05GY53 XANOR	05gy53	xanthomonas
711	15	57.7	122	2	06SRX5 CVHSA	06srx5	sars	corona	784	15	57.7	139	2	04MNG4 BACCB	04mng4	bacillus ce
712	15	57.7	122	2	06SRY0 CVHSA	06sry0	sars	corona	785	15	57.7	139	2	06HFI1 BACHK	06hfi1	bacillus th
713	15	57.7	122	2	06SRV5 CVHSA	06sry5	sars	corona	786	15	57.7	139	2	0733N2 BACCI	0733n2	bacillus ce
714	15	57.7	122	2	06SRZ0 CVHSA	06srz0	sars	corona	787	15	57.7	139	2	081AG4 BACCR	081ag4	bacillus ce
715	15	57.7	122	2	06VA72 CVHSA	06va72	sars	corona	788	15	57.7	139	2	097D28 CLOAB	097d28	clostridium
716	15	57.7	122	2	06VA83 CVHSA	06va83	sars	corona	789	15	57.7	139	2	081Y90 BACAN	081y90	bacillus an
717	15	57.7	122	2	06VA94 CVHSA	06va94	sars	corona	790	15	57.7	139	2	0637J9 BACCC	0637j9	bacillus ce
718	15	57.7	122	2	04JDP8 CVHSA	04jdp8	sars	corona	791	15	57.7	139	2	0699G3 9H1V1	0699g3	human immun
719	15	57.7	122	2	04JDR1 CVHSA	04jdr1	sars	corona	792	15	57.7	140	1	LY6H HUMAN	ly6h	human
720	15	57.7	122	2	06C149 XENIA	06c149	xenopus	lae	793	15	57.7	140	2	061AX0 HUMAN	061ax0	homo sapien
721	15	57.7	123	2	09JQZ2 NEIMA	09jqz2	neisseria	m	794	15	57.7	140	2	04V3U1 DROME	04v3u1	drosophila
722	15	57.7	123	2	07DDH5 NEIMB	07ddh5	neisseria	m	795	15	57.7	140	2	06GZ25 ORYSA	06gz25	oryza sativ
723	15	57.7	123	2	06Y908 9H1V1	06y908	human	immun	796	15	57.7	140	2	08AR11 9H1V1	08ar11	human immun
724	15	57.7	124	2	09PKS3 CHLAMU	09pks3	chlamydia	m	797	15	57.7	141	2	06GKZ6 DROME	06gkz6	drosophila
725	15	57.7	124	2	07SR41 9H1V1	07sr41	human	immun	798	15	57.7	141	2	0664U8 VERPS	0664u8	versinia ps
726	15	57.7	125	2	06KDD2 ECOLI	06kdd2	escherichia		799	15	57.7	141	2	08ZJ85 VERPE	08zj85	versinia pe
727	15	57.7	125	2	04KFP0 PSBFS	04kfp0	pseudomonas		800	15	57.7	142	2	09YF15 DROME	09yf15	drosophila
728	15	57.7	125	2	08F1Z3 ECOL6	08f1z3	escherichia		801	15	57.7	143	2	06ETJ0 HUMAN	06etj0	homo sapien
729	15	57.7	125	2	04TEY5 TETNG	04tey5	tetradon n		802	15	57.7	143	2	09S406 HORVD	09s406	hoddeum vul
730	15	57.7	125	2	06RUW8 9H1V1	06ruw8	human	immun	803	15	57.7	143	2	09ACU2 STREPMC	09acu2	streptococ
731	15	57.7	126	2	0417X6 GIBZB	0417x6	giardiera		804	15	57.7	143	2	06MH33 BDBEA	06mh33	bdellovibri
732	15	57.7	126	2	07R635 GIARA	07r635	giardiera	lam	805	15	57.7	144	2	04F8J1 EUCGL	04f8j1	eucalyptus
733	15	57.7	126	2	088C21 PSBPK	088c21	pseudomonas		806	15	57.7	144	2	08CZ18 VERPE	08cz18	versinia pe
734	15	57.7	127	2	04ZL69 PSBSY	04zl69	pseudomonas		807	15	57.7	145	2	08WQ22 LOCMI	08wq22	locusta mg
735	15	57.7	127	2	069962 9H1V1	069962	human	immun	808	15	57.7	145	2	09VZ15 DROME	09vz15	drosophila
736	15	57.7	128	1	KRA24 HUMAN	kra24	homo sapien		809	15	57.7	145	2	06K6Z6 ORYSA	06k6z6	oryza sativ
737	15	57.7	128	2	08LBO3 ARATH	08lbo3	arabidopsis		810	15	57.7	146	2	05K4F7 SCGGR	05k4f7	schistosom
738	15	57.7	128	2	04ZU08 PSBSY	04zu08	pseudomonas		811	15	57.7	146	2	07XNZ5 ORYSA	07xnz5	oryza sativ
739	15	57.7	129	2	05D9M3 SCHUA	05d9m3	schistosoma		812	15	57.7	146	2	08AXW7 MICCO	08axw7	micritus co
740	15	57.7	129	2	06IKR4 DROME	06ikr4	drosophila		813	15	57.7	146	2	04RFX5 TETNG	04rfx5	tetradon n
741	15	57.7	129	2	05OKR5 LYNST	05okr5	lymaea	sta	814	15	57.7	147	2	05Y2U3 9H1V1	05y2u3	human immun
742	15	57.7	129	2	04IVL5 AZOVI	04ivl5	azocobacter		815	15	57.7	148	2	06ZMW2 HUMAN	06zmw2	homo sapien
743	15	57.7	130	2	054OR7 DICDI	054or7	dictyosteli		816	15	57.7	148	2	05TC71 HUMAN	05tc71	homo sapien
744	15	57.7	130	2	06THB0 DROME	06thb0	drosophila		817	15	57.7	148	2	04RL35 TETNG	04rl35	tetradon n
745	15	57.7	131	2	09S237 STRCO	09s237	streptomyce		818	15	57.7	149	2	08UAP1 AGRT5	08uap1	agrobacteri
746	15	57.7	131	2	06CZ68 ERWCT	06cz68	erythra	car	819	15	57.7	149	2	0898V0 CIOTE	0898v0	ciostriidum
747	15	57.7	132	2	05S2H5 CRYNE	05s2h5	cryptococcu		820	15	57.7	150	2	05AC28 CANAL	05ac28	candida alb
748	15	57.7	132	2	05D995 SCHUA	05d995	schistosoma		821	15	57.7	150	2	06ICF4 CAEBR	06icf4	caenothabdi
749	15	57.7	132	2	09AN59 BRAJA	09an59	bradyrhizob		822	15	57.7	150	2	06LIDZ5 PARTE	06lidz5	parametium
750	15	57.7	132	2	07SOV8 9H1V1	07sov8	human	immun	823	15	57.7	151	2	09BHM5 CABEL	09bhm5	caenorhabdi
751	15	57.7	133	2	04TZU4 TETNG	04tzu4	tetradon n		824	15	57.7	152	2	08TIR4 METAC	08tir4	methanosarc
752	15	57.7	133	2	08U0U5 9H1V1	08u0u5	human	immun	825	15	57.7	152	2	07PXZ2 ANOGA	07pxz2	anopheles g
753	15	57.7	134	2	024959 GIARA	024959	giardia	lam	826	15	57.7	153	2	080876 BSMV	080876	barley etci
754	15	57.7	134	2	024988 GIARA	024988	giardia	lam	827	15	57.7	153	2	094178 ORYSA	094178	oryza sativ
755	15	57.7	134	2	04V6M8 DROME	04v6m8	drosophila		828	15	57.7	154	2	05TPK5 ANOGA	05tpk5	anopheles g
756	15	57.7	134	2	06RUX2 9H1V1	06rux2	human	immun	829	15	57.7	155	2	09SKW0 BOVIN	09skw0	boe taurus
757	15	57.7	135	1	YDQA SCHPO	ydqa	schizosacch		830	15	57.7	155	2	08GSR9 WHEAT	08gsr9	triticum ae
758	15	57.7	135	2	05DBE6 SCHUA	05dbe6	schistosoma		831	15	57.7	155	2	069536 9BETA	069536	human herpe
759	15	57.7	135	2	04QIR1 LEIMA	04qir1	leishmania		832	15	57.7	155	2	01Z317 9H1V1	01z317	human immun
760	15	57.7	135	2	084YV6 ORYSA	084yv6	oryza sativ		833	15	57.7	156	2	08QIM4 9NUCL	08qim4	manebstra co
761	15	57.7	135	2	08H6Y5 PHYTN	08h6y5	phytophthor		834	15	57.7	157	1	CI0255 HUMAN	ci0255	homo sapien

961 15 57.7 207 2 Q4VE68 DROME
962 15 57.7 207 2 Q70016 SHIDY
963 15 57.7 207 2 Q8FCJ3 ECOL6
964 15 57.7 207 2 Q8XSN4 ECOS7
965 15 57.7 208 2 Q6XHV9 DROYA
966 15 57.7 208 2 Q82069 SOLTU
967 15 57.7 208 2 Q4R8J3 EUGCL
968 15 57.7 208 2 Q9AHJ9 YLACO
969 15 57.7 209 2 Q5SGM2 DICI1
990 15 57.7 210 2 Q7KML7 DICI1
991 15 57.7 210 2 Q7XU70 ORYSA
992 15 57.7 210 2 Q93CB9 MYCBA
993 15 57.7 210 2 Q39318 SHIV1
994 15 57.7 211 1 G12 ANOGA
995 15 57.7 211 2 Q7Q5V5 ANOGA
996 15 57.7 211 2 Q50SC9 ENTHI
997 15 57.7 211 2 Q6WES3 PTLEL
998 15 57.7 212 2 Q50PE5 ENTHI
999 15 57.7 212 2 Q9XV06 CAEEL
1000 15 57.7 212 2 Q4Q6X5 LETMA

ALIGNMENTS

RESULT 1
ID Q96KM3 HUMAN PRELIMINARY; PRT; 213 AA.

AC Q96KM3;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Aberrant WW domain-containing oxidoreductase.
GN Name=WWOX;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OX NCB1_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=2145735; PubMed=11572989; DOI=10.1073/pnas.191175896;
RA Paige A.J.W., Taylor K.J., Taylor C., Hillier S.G., Farrington S.,
RT Scott D., Porteous D.J., Smyth J.F., Gabra H., Watson J.E.V.,
RT "WWOX: a candidate tumor suppressor gene involved in multiple tumor
types".
RT Proc. Natl. Acad. Sci. U.S.A. 98:11417-11422(2001).
RL EMBL; AF325432; AAL05451.1; -; Genomic DNA.
DR EMBL; AF325423; AAL05451.1; JOINED; Genomic DNA.
DR EMBL; AF325424; AAL05451.1; JOINED; Genomic DNA.
DR EMBL; AF325426; AAL05451.1; JOINED; Genomic DNA.
DR EMBL; AF325433; AAL05451.1; JOINED; Genomic DNA.
DR EMBL; AF325425; AAL05451.1; JOINED; Genomic DNA.
DR HSSP; Q13526; IPTN.
DR Ensembl; ENSG0000186153; Homo sapiens.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00397; WW; 2.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS01159; WW_DOMAIN 1; 2.
DR PROSITE; PS50020; WW_DOMAIN 2; 2.
SQ SEQUENCE 213 AA; 23866 MW; A21054FF8214CC7C CRC64;

Query Match 69.2%; Score 18; DB 2; Length 213;
Best Local Similarity 20.0%; Pred. No. 0.43;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 184 CTTSTAAC 193

RESULT 2
Q586F4_9TRYP

ID Q586F4_9TRYP PRELIMINARY; PRT; 263 AA.

AC Q586F4;
DT 10-MAY-2005 (TRENBLrel. 30, Created)
DT 10-MAY-2005 (TRENBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Rp927.6.4820;
OS Trypanosoma brucei.
OC Trypanosoma; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OC Eukaryota; Euklenozoa;
OX NCB1_TaxID=5691;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUtatio.1;
RA Chedin E., Blandin G., Bartholomeu D., Caler E., Haas B., Hannick L.,
RA Johnson J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,
RA Shallow J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
RA Khalak H.G., Salzberg S., Simpson A.J., Tallon L., Van Aken S.,
RA Wainless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUtatio.1;
RA El-Sayed N.M., Khalak H., Adams M.D.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUtatio.1;
RA Haas B., Blandin G., El-Sayed N.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC008146; AAX80300.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 263 AA; 28873 MW; F61DF5AF910531AB CRC64;

Query Match 69.2%; Score 18; DB 2; Length 263;
Best Local Similarity 20.0%; Pred. No. 0.46;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 95 CATTSSSAC 104

RESULT 3

ID Q7PRO7 ANOGA PRELIMINARY; PRT; 269 AA.

AC Q7PRO7;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE ENSANGP0000001657 (Fragment).
GN ORFNames=ENSANGG0000001387;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCB1_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation".
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.

DR EMBL; AAB01008847; EAA06779.3; -; Genomic DNA.
FT NON_TER 1 1
FT 269 269

SEQUENCE 269 AA; 24228 MW; 25BBP71FD71F1F2 CRC64;

Query Match Score 18; DB 2; Length 269;

Best Local Similarity 20.0%; Pred. No. 0.47; Mismatches 8; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

1 CXXXXXXXXXXC 10

149 CTTSSSTTSC 158

RESULT 4

OSTX10 ANOGA PRELIMINARY; PRT; 281 AA.

AC 05TX10; 01-FEB-2005 (TREMBlrel. 29, Created)

DT 01-FEB-2005 (TREMBlrel. 29, Last sequence update)

DE ENSANGP0000027659 (Fragment).

GN ORFNames=ENSANG00000007023;

OS Anopheles gambiae str. PEPT.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;

OC Anophelinae; Anopheles.

NCBI_TaxID=180454;

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;

RG The Anopheles gambiae Sequence Committee;

RT "Anopheles gambiae re-annotation."

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

CC -1- CAUTION: The sequence shown here is derived from an

EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

DR EMBL; AAB01008807; EAL41760.1; -; Genomic_DNA.

FT NON_TER 281 281

FT SEQUENCE 281 AA; 25341 MW; 0462E1169F8642B CRC64;

Query Match Score 18; DB 2; Length 281;

Best Local Similarity 20.0%; Pred. No. 0.48; Mismatches 8; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

1 CXXXXXXXXXXC 10

107 CTTAASSTAC 116

RESULT 5

YMS8 YEAST STANDARD; PRT; 313 AA.

AC 003655; 01-NOV-1997 (rel. 35, Created)

DT 01-NOV-1997 (rel. 35, Last sequence update)

DE Hypothetical 35.0 kDa protein in PK2-HF1 intergenic region.

GN OrderedAccession=YMR206W; ORFNames=YMR325.07;

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycace.

NCBI_TaxID=4932;

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=S288C / AB972;

RA MEDLINE=97313268; PubMed=9169872;

RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,

RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,

RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,

RA Rice P., Skelton J., Walsh S.V., Whitehead S., Barrell B.G.;

RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome

III."

RL Nature 387:90-93 (1997).

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC -----

DR EMBL; 248755; CAA88648.1; -; Genomic_DNA.

DR PIR; S59448; S59448.

DR GenBank; 142881; -.

DR Ensembl; YMR206W; Saccharomyces cerevisiae.

DR SGD; S000004819; YMR206W.

KM Complete proteome; Hypothetical protein.

FT COMEBIAS 3 6 Poly-Ser.

FT COMEBIAS 146 149 Poly-Gln.

FT COMEBIAS 246 252 Poly-Ser.

FT SEQUENCE 313 AA; 35018 MW; 9D92BFD8982577F0 CRC64;

Query Match Score 18; DB 1; Length 313;

Best Local Similarity 20.0%; Pred. No. 0.5; Mismatches 8; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

1 CXXXXXXXXXXC 10

245 CTTSSSTTSC 254

RESULT 6

Q6ZDR4 ORYSA PRELIMINARY; PRT; 357 AA.

AC Q6ZDR4; 05-JUL-2004 (TREMBlrel. 27, Created)

DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)

DE 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

GN Name=P0481F05.20;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

NCBI_TaxID=39947;

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, PAC

clone: P0481F05."

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).

DR EMBL; AP004376; BAD09322.1; -; Genomic_DNA.

DR HSSP; Q03237; IASU.

DR Gramene; Q6ZDR4; -.

DR GO; GO:0005634; C:nucleus; IEA.

DR InterPro; IPR012287; Homeodomain-rel.

DR Pfam; PF00249; MYB_DNA_Bd.

DR SMART; SM00717; SANT; 2.

DR PROSITE; PS00037; MYB_1; UNKNOWN_1.

DR PROSITE; PS00034; MYB_2; 1.

DR PROSITE; PS00090; MYB_3; 2.

DR Nuclear protein; Repeat.

SK SEQUENCE 357 AA; 36754 MW; 3FAF56860D3BAE22 CRC64;

Query Match Score 18; DB 2; Length 357;

Best Local Similarity 20.0%; Pred. No. 0.52; Mismatches 8; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

1 CXXXXXXXXXXC 10

DB 322 CSATASASC 331

RESULT 7

ID O7QJL0 ANOGA PRELIMINARY; PRT; 362 AA.

AC O7QJL0;

DT 01-MAR-2004 (TREMBlrel. 26, Created)

DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)

DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

DE ENSANGP0000009383 (Fragment).

GN ORFNames=ENSANGG0000007023;

OS Anopheles gambiae str. PEST.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;

OC Anophelinae; Anopheles.

OX NCBI_TaxID=180454;

XX [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;

RG The Anopheles gambiae Sequence Committee;

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

CC -1- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

DR EMBL; AAAB01008807; EAA04378.2; -; Genomic_DNA.

FT NON_TER 1

SO SEQUENCE 362 AA; 32490 MW; 384376DDFA31BEC0 CRC64;

Query Match 69.2%; Score 18; DB 2; Length 362;

Best Local Similarity 20.0%; Pred. No. 0.52;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10

DB 187 CTAASSTAC 196

RESULT 8

Q9RV87 DEIRA PRELIMINARY; PRT; 480 AA.

ID Q9RV87;

AC Q9RV87;

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Hypothetical protein DR1142.

GN OrderedLocNames=DR1142;

OS Deinococcus radiodurans.

OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;

OC Deinococcaceae; Deinococcus.

OX NCBI_TaxID=1299;

XX [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=1 / ATCC 13939 / DSM 20539 / NCIB 9279;

RX MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,

RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,

RA Vamathevan J.J., Lam P., McDonald L.A., Ueberback T.R., Zalewski C.,

RA Makarov K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,

RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,

RA Fraser C.M.;

RT "Genome sequence of the radioresistant bacterium Deinococcus

RT radiodurans R1.";

RL Science 286:1571-1577(1999).

DR EMBL; AE001963; AAP10716.1; -; Genomic_DNA.

DR PIR; E75433; E75433.

DR TIGR: DR1142; -;

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 480 AA; 49364 MW; 1B962ADAC24AB CRC64;

Query Match 69.2%; Score 18; DB 2; Length 480;

Best Local Similarity 20.0%; Pred. No. 0.58;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10

DB 464 CASAAATAC 473

RESULT 9

ID TWCC1_HUMAN STANDARD; PRT; 653 AA.

AC O94876; Q68E06; Q81XM8;

DT 16-OCT-2001 (Rel. 40, Created)

DT 10-MAY-2005 (Rel. 47, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE Transmembrane and coiled-coil domains protein 1.

GN Name=TWCC1; Synonyms=K1AA0779;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;

OC Homo.

OX NCBI_TaxID=9606;

XX [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RC TISSUE=Skin;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McSwan P.J., McKernan K.D., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Gardia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 15-653.

RC TISSUE=Endometrial tumor;

RG The German cDNA consortium;

RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.

RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 334-653.

RC TISSUE=Brain;

RX MEDLINE=99087487; PubMed=9872452;

RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,

RA Tanaka A., Kotani H., Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. XI.

RT The complete sequences of 100 new cDNA clones from brain which code

RT for large proteins in vitro.";

RL DNA Res. 5:277-286(1998).

XX -1- SIMILARITY: Belongs to the TEX28 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

```

CC -----
DR EMBL; AB018322; BAA34499.1; -; mRNA.
DR EMBL; BC039859; AAH39859.1; -; mRNA.
DR EMBL; CR749206; CAH18064.2; -; mRNA.
DR Ensemble; ENSG00000172765; Homo sapiens.
DR HGNC; HGNC:29116; TMCC1.
KW Coiled coil; Transmembrane.
FT TRANSMEM 592 612 Potential.
FT TRANSMEM 625 645 Potential.
FT COILED 228 313 Potential.
FT COILED 458 576 Potential.
SQ SEQUENCE 653 AA; 72053 MW; 18B07D171B874205 CRC64;

Query Match
Best Local Similarity 69.2%; Score 18; DB 1; Length 653;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
Db 173 CAAAAAAC 182

RESULT 10
ID 05H317_XANOR PRELIMINARY; PRT; 1005 AA.
AC 05H317;
DT 10-MAY-2005 (TReMBLrel. 30, Created)
DT 10-MAY-2005 (TReMBLrel. 30, Last sequence update)
DE Transcriptional regulator.
GN Name=acok; OrderedlocusNames=XO01230;
OS Xanthomonas oryzae (pv. oryzae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxId=64187;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KACC10331 / KX085;
RX PubMed=15673718; DOI=10.1093/nar/gk1206;
RA Lee B.-M., Park Y.-J., Park D.-S., Kang H.-W., Kim J.-G., Song E.-S.,
RA Park I.-C., Yoon U.-H., Hahn J.-H., Koo B.-S., Lee G.-B., Kim H.,
RA Park H.-S., Yoon K.-O., Kim J.-H., Jung C.-H., Koh N.-H., Seo J.-S.,
RA Go S.-J.;
RT "The genome sequence of Xanthomonas oryzae pathovar oryzae KACC10331,
RT the bacterial blight pathogen of rice.";
RL Nucleic Acids Res. 33:577-586(2005).
DR EMBL; AE013598; AAW74484.1; -; Genomic DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:001711; P:nucleoside-triphosphatase activity; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR011717; TPR_4.
DR Pfam; PF00196; GeTB; 1.
DR Pfam; PF00515; TPR_1; 2.
DR Pfam; PF07721; TPR_4; 1.
DR PRINTS; PR00038; HTHLTKR.
DR ProDom; PD000307; HTH_LuxR; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00421; HTH_LuxR; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS00622; HTH_LuxR_FAMILY; 1.
DR PROSITE; PS00293; TPR_REGION; 1.
KW Complete proteome.
SQ SEQUENCE 1005 AA; 110791 MW; 662961379FD37B61 CRC64;

Query Match
Best Local Similarity 69.2%; Score 18; DB 2; Length 1005;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

Qy 1 CXXXXXXXXXC 10
Db 386 CASATTAAC 395

RESULT 11
ID 040CX7_LEIMA PRELIMINARY; PRT; 1061 AA.
AC 040CX7;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Kinesin heavy chain, putative.
GN ORFNames=LmjF20_0640;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxId=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005259; CAJ03791.1; -; Genomic DNA.
SQ SEQUENCE 1061 AA; 113930 MW; 2C69385D06371DCC CRC64;

Query Match
Best Local Similarity 69.2%; Score 18; DB 2; Length 1061;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
Db 192 CAATATSTSC 201

RESULT 12
ID 08LPD6_HORVU PRELIMINARY; PRT; 52 AA.
AC 08LPD6;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Putative glutathione-S-transferase (Fragment).
GN Name=bar2;
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OX NCBI_TaxId=4513;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Baldwin A.;
RT "The effect of thiocarbamate herbicides and the safener dichloromid in
RT barley and wild oats.";
RL Thesis (2001), Department of Cardiff School of Biosciences, Cardiff
RL university, Cardiff, United Kingdom.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Baldwin A., Harwood J.L., Machray G.C., Francis D., Rogers H.J.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ419772; CAD11963.1; -; mRNA.
DR HSSP; P12653; IAXD.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002197; HTH_Fls.
DR PRINTS; PR01590; HTHFIS.
KW Transferase.
SQ SEQUENCE 52 AA; 5750 MW; 7392BB72A3C70B9B CRC64;

FT NON_TER 1 1
FT NON_TER 52 52
SQ SEQUENCE 52 AA; 5750 MW; 7392BB72A3C70B9B CRC64;

```

Query Match 65.4%; Score 17; DB 2; Length 52;
Best Local Similarity 20.0%; Pred. No. 1.3;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
|
Db 18 CTSASSTRSC 27

RESULT 13

Q5DCN3_SCHJA PRELIMINARY; PRT; 122 AA.
ID Q5DCN3_SCHJA
AC Q5DCN3;
DT 10-MAY-2005 (TREMBLrel. 30, Created)
DT 10-MAY-2005 (TREMBLrel. 30, last sequence update)
DT 10-MAY-2005 (TREMBLrel. 30, last annotation update)
DE Hypothetical protein.
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OC NCBI_TaxId=6182;
OX NCBI_TaxId=6182;
RN NUCLEOTIDE SEQUENCE.
RP Han Z.;
RA "The full-length cDNA sequences of Schistosoma japonicum genes."
RT Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY814691; FAY814691; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 122 AA; 13028 MW; 960E8615986958DE CRC64;

Query Match 65.4%; Score 17; DB 2; Length 122;
Best Local Similarity 20.0%; Pred. No. 1.8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
|
Db 76 CTSASSTRSC 85

RESULT 14

Q96LJ4_HUMAN PRELIMINARY; PRT; 128 AA.
ID Q96LJ4_HUMAN
AC Q96LJ4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, last annotation update)
DE Hypothetical protein FLJ25437.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
OC Homo.
OC NCBI_TaxId=9606;
OX NCBI_TaxId=9606;
RN NUCLEOTIDE SEQUENCE.
RP TissueTestis;
RA Kawakami B., Sugiyama A., Takemoto M., Suzuki Y., Hata H.,
RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,
RA Irie K., Otsuki T., Sato H., Nishikawa T., Nagai K., Isogai T.,
RA Sugano S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK058166; BAB71697.1; -; mRNA.
DR Ensembl; ENSG00000178093; Homo sapiens.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; Kinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Kinase; Nucleotide-binding; Transferase;
SQ SEQUENCE 128 AA; 14116 MW; D30AC8C668621B0C CRC64;

Query Match 65.4%; Score 17; DB 2; Length 128;
Best Local Similarity 20.0%; Pred. No. 1.8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
|
Db 118 CTAATSSAKTC 127

RESULT 15

O6ZKM7_ORYSA PRELIMINARY; PRT; 139 AA.
ID O6ZKM7_ORYSA
AC O6ZKM7;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, last annotation update)
DE Hypothetical protein OJ1118_A06.7-1.
GN Name=OJ1118_A06.7-1;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxId=39947;
OX NCBI_TaxId=39947;
RN NUCLEOTIDE SEQUENCE.
RP Sasaki T., Matsunoto T., Yamamoto K.;
RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
RT clone:OJ1118_A06.7";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003873; BAD08805.1; -; Genomic DNA.
DR Gramene; O6ZKM7; -;
KW Hypothetical protein.
SQ SEQUENCE 139 AA; 14109 MW; E281360370FB63C9 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 139;
Best Local Similarity 20.0%; Pred. No. 1.9;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
|
Db 77 CAAITSSSAC 86

RESULT 16

O6H482_ORYSA PRELIMINARY; PRT; 155 AA.
ID O6H482_ORYSA
AC O6H482;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, last sequence update)
DT 10-MAY-2005 (TREMBLrel. 30, last annotation update)
DE Hypothetical protein OJ1119_C03.6 (Hypothetical protein
DE B1012G11.50).
GN Name=OJ1119_C03.6; Synonyms=B1012G11.50;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxId=39947;
OX NCBI_TaxId=39947;
RN NUCLEOTIDE SEQUENCE.
RP Sasaki T., Matsunoto T., Hattori M., Sasaki Y., Katayose Y.;
RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, BAC
RT clone:OJ1119_C03.6";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005903; BAD38424.1; -; Genomic DNA.
SQ SEQUENCE 155 AA; 14116 MW; D30AC8C668621B0C CRC64;

Query Match 65.4%; Score 17; DB 2; Length 128;
Best Local Similarity 20.0%; Pred. No. 1.8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
|
Db 18 CTSASSTRSC 27

DR Gramene; G6H482; -;
KW Hypothetical protein.
SQ SEQUENCE 155 AA; 16233 MW; 02D836E58C7B52 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 155;
Best Local Similarity 20.0%; Pred. No. 2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXX 10
DB 25 CTTSAADC 34

RESULT 17

0657M0 ORYSA PRELIMINARY; PRT; 160 AA.

AC 0657M0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein P0468H06.13 (Hypothetical protein P0468H05.38).
CN Name=P0468H06.13; Synonyms=P0468H05.38;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
DR EMBL; CAAB0100659; CAP92211.1; -; Genomic DNA.
SQ SEQUENCE 160 AA; 17094 MW; 6167384F2A5E51B CRC64;

NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nishimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizuyashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shinokawa T., Shimura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhang H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yang M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003075; BAD44997.1; -; Genomic DNA.
DR EMBL; AP003267; BAD87274.1; -; Genomic DNA.
DR Gramene; O657M0; -;
KW Hypothetical protein.
SQ SEQUENCE 160 AA; 17630 MW; C3A5ABDE5CF75DA9 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 160;
Best Local Similarity 20.0%; Pred. No. 2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXX 10
DB 105 CSGATSTASC 114

RESULT 18

04T4H1 TETNG PRELIMINARY; PRT; 160 AA.

AC 04T4H1;
DT 13-BEP-2005 (TrEMBLrel. 31, Created)
DT 13-BEP-2005 (TrEMBLrel. 31, Last annotation update)
DT 13-BEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCARF659, whole genome shotgun sequence.
CN ORFName=GSTENG00007326001;
OS Tetradon nigrovittidis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetradontiformes;
OC Tetradontidae; Tetradontidae; Tetradon.
OX NCBI_TaxId=99883;
RN [1]

NUCLEOTIDE SEQUENCE.

RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Maucel E., Bouneau L., Fischer C., Ozouf-Costat C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Catellico L., Poulain J., De Bernardis V.,
RA Cruaud C., Duprat S., Broctier P., Coutanceau J.P., Guzy J.,
RA Parra G., Lardier S., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Queller F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Croallins H.;
RT "Genome duplication in the teleost fish Tetradon nigrovittidis reveals the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]

NUCLEOTIDE SEQUENCE.

RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB0100659; CAP92211.1; -; Genomic DNA.
SQ SEQUENCE 160 AA; 17094 MW; 6167384F2A5E51B CRC64;

Query Match 65.4%; Score 17; DB 2; Length 160;
Best Local Similarity 20.0%; Pred. No. 2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXX 10
DB 115 CTTASSTSC 124

RESULT 19

0825B8 STRAW PRELIMINARY; PRT; 219 AA.

AC 0825B8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative tetr-family transcriptional regulator.
GN Ordered locus names=SAV7510;
OS Streptomyces avermitilis;
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
OC Streptomycetaceae; Streptomycetaceae; Streptomyces.
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
DR EMBL; CAAB0100659; CAP92211.1; -; Genomic DNA.
SQ SEQUENCE 160 AA; 17094 MW; 6167384F2A5E51B CRC64;

Query Match 65.4%; Score 17; DB 2; Length 160;
Best Local Similarity 20.0%; Pred. No. 2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXX 10
DB 115 CTTASSTSC 124

RESULT 19

0825B8 STRAW PRELIMINARY; PRT; 219 AA.

AC 0825B8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative tetr-family transcriptional regulator.
GN Ordered locus names=SAV7510;
OS Streptomyces avermitilis;
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
OC Streptomycetaceae; Streptomycetaceae; Streptomyces.
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
DR EMBL; CAAB0100659; CAP92211.1; -; Genomic DNA.
SQ SEQUENCE 160 AA; 17094 MW; 6167384F2A5E51B CRC64;

```
DR GO: 0003700; F:transcription factor activity; IEA.
DR GO: 0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR012287; Homeodomain-rel.
DR InterPro: IPR01647; HTH_Tetr.
DR Pfam: PF00440; Tetr_N; 1.
DR PROSITE: PS50977; HTH_TERR_2; 1.
KW Complete proteome.
SQ SEQUENCE 219 AA; 24066 MW; 5C59ADP39430F304 CRC64;

Query Match
Best Local Similarity 65.4%; Score 17; DB 2; Length 219;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
DB 151 CAEAAAAAAC 160

RESULT 20
O6152_9VIRU PRELIMINARY; PRT; 220 AA.
ID O6152_9VIRU
AC O6152;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Coat protein.
OS Cymbidium mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flexiviridae;
OC Potexvirus.
OX NCBI_TaxId=12178;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Korean isolate;
RX MEDLINE=95278762; PubMed=7758973; DOI=10.1016/0378-1119(95)00105-F;
RA Ryu K.H., Yoon K.B., Park W.M.;
RT "Nucleotide sequence of coat protein gene of cymbidium mosaic
RT potexvirus genomic RNA, the Korean isolate.";
RL Gene 156:303-304(1995).
DR EMBL: X81051; CAAS6941.1; -; Genomic_RNA.
DR PIR: JC4082; JC4082.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR000052; P1vtr_coat.
DR Pfam: PF00286; Vtrub_P-coat_1.
DR PRINTS: PR00232; POTCARLCOAT.
DR ProDom: PD00603; P1vtr_coat; 1.
KW Capsid protein.
SQ SEQUENCE 220 AA; 23719 MW; 1F40C6E7CDCCF3B8 CRC64;

Query Match
Best Local Similarity 65.4%; Score 17; DB 2; Length 220;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
DB 166 CSAATITATC 175

RESULT 21
O6YTS2_ORYSA PRELIMINARY; PRT; 256 AA.
ID O6YTS2_ORYSA
AC O6YTS2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein P0419H09.29.
GN Name=P0419H09.29;
OS Oryza sativa (japonica cultivar-group). Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Streptophyta; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxId=39947;
RN [1]
```

```
RP NUCLEOTIDE SEQUENCE.
RA Saeki T., Matsunoto T., Katayose Y.;
RL Submitted (NOV-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AP005918; BAD10719.1; -; Genomic_DNA.
DR Gramene: O6YTS2; -.
KW Hypothetical protein.
SQ SEQUENCE 256 AA; 26591 MW; 63BFCAF53D2D40EA CRC64;

Query Match
Best Local Similarity 65.4%; Score 17; DB 2; Length 256;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
DB 203 CSASASARC 212

RESULT 22
O65023_ORYSA PRELIMINARY; PRT; 260 AA.
ID O65023_ORYSA
AC O65023;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE DnaU-like protein (Putative heat shock protein).
GN Name=OSUNB0027B08.12; Synonym=OSUNB0078D06.16;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxId=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tstirln T., Kim M.M., Bera J.V., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Hsieh J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Uteback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (FEB-2003) to the EMBL/Genbank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (MAY-2003) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: Have a continuous role in plant development probably in
CC the structural organization of compartments (By similarity).
DR EMBL: AC133778; MA039854.1; -; Genomic_DNA.
DR EMBL: AC133339; AAP4230.1; -; Genomic_DNA.
DR HSSP: P25685; 1HDJ.
DR Gramene: O65023; -.
DR GO: GO:0031072; F:heat shock protein binding; IEA.
DR GO: GO:0051082; F:unfolded protein binding; IEA.
DR GO: GO:0006457; P:protein folding; IEA.
DR GO: GO:0006986; P:response to unfolded protein; IEA.
DR InterPro: IPR001623; DnaU_N.
DR InterPro: IPR003095; Hsp_DnaU.
DR Pfam: PF00226; DnaU; 1.
DR PRINTS: SM00271; DnaU; 1.
DR SMART: SM00636; DnaU_1; 1.
DR PROSITE: PS50076; DnaU_2; 1.
KW Chaperone; Heat shock.
SQ SEQUENCE 260 AA; 28639 MW; DDC645DAF9BA1ED CRC64;

Query Match
Best Local Similarity 65.4%; Score 17; DB 2; Length 260;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
```


Db 19 CAAAFSAC 28

RESULT 23

Q57E4_ANOGA PRELIMINARY; PRT; 304 AA.
AC Q57E4;
DT 01-FEB-2005 (TREMBlrel. 29, Created)
DT 01-FEB-2005 (TREMBlrel. 29, Last sequence update)
DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
DE ENSANGP000025853 (Fragment).
GN ORFNames=ENSANG0000024781;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxId=180454;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PEST;
RC The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01006839; EAL41365.1; -; Genomic_DNA.
FT NON_TER 1 304
FT SEQUENCE 304 AA; 29530 MW; 1F55F0FB46B16B9 CRC64;
SQ

Query Match 65.4%; Score 17; DB 2; Length 304;
Best Local Similarity 20.0%; Pred. No. 2.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10
Db 20 CSSTSTVC 29

RESULT 24

Q582B8_9TRYP PRELIMINARY; PRT; 344 AA.
AC Q582B8;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=TP927.5.2480;
OS Trypanosoma brucei.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxId=5691;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=GUTat10.1;
RC Ghedin E., Blandin G., Bartholomew D., Caler E., Haas B., Hannick L.,
RA Shalom J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,
RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
RA Khalak H.G., Salzberg S., Simpson A.U., Tallon L., Van Aken S.,
RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA El-Sayed N.M., Khalak H., Adams M.D.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
RN [3]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Haas B., Blandin G., El-Sayed N.;
RL Submitted (Apr-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC091655; AAX80451.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 344 AA; 37545 MW; A3925C9B2A48BD53 CRC64;
KM

Query Match 65.4%; Score 17; DB 2; Length 344;
Best Local Similarity 20.0%; Pred. No. 2.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10
Db 251 CSSTASGSC 260

RESULT 25

Q57Z20_9TRYP PRELIMINARY; PRT; 361 AA.
AC Q57Z20;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=TP927.5.4290;
OS Trypanosoma brucei.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxId=5691;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=GUTat10.1;
RC Ghedin E., Blandin G., Bartholomew D., Caler E., Haas B., Hannick L.,
RA Shalom J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,
RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
RA Khalak H.G., Salzberg S., Simpson A.U., Tallon L., Van Aken S.,
RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA El-Sayed N.M., Khalak H., Adams M.D.;
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Haas B., Blandin G., El-Sayed N.;
RL Submitted (Apr-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116668; AAX80579.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 361 AA; 40083 MW; 95BC9BDCB3B9895 CRC64;
KM

Query Match 65.4%; Score 17; DB 2; Length 361;
Best Local Similarity 20.0%; Pred. No. 2.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10
Db 205 CSSALSSSC 214

RESULT 26

Q4YW87_PLABE PRELIMINARY; PRT; 364 AA.
AC Q4YW87;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
GN ORFNames=PB105184.00.0;
OS Plasmodium berghei.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxId=5821;

Db 17 CAAATKAC 26

RESULT 29

Q7PY13_ANOGA PRELIMINARY; PRT; 388 AA.

AC Q7PY13; 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE ENSANGP0000018378 (Pragment).
 GN ORFNames=ENSANG0000015889;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
 CC Anophelinae; Anophelinae;
 CC NCBI_Taxid=180454;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PEST;
 RA Anopheles Genome Sequencing Consortium;
 RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.

DR EMBL; AAAB01008987; EAA01136.1; -; Genomic_DNA.
 DR InterPro; IPR000195; RabGAP_TBC.
 DR Pfam; PF00566; TBC; 1.
 DR PROSITE; PS50086; TBC_RABGAP; 1.
 FT NON TER 388
 SQ SEQUENCE 388 AA; 44534 MW; A3EEC567D458CCFF CRC64;

Query Match 65.4%; Score 17; DB 2; Length 388;
 Best Local Similarity 20.0%; Pred. No. 2.8;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXC 10
 DB 38 CSTSAFTC 47

RESULT 30
 Q21081_CAEEL PRELIMINARY; PRT; 389 AA.

AC Q21081; 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein K01C8.2.
 GN ORFNames=K01C8.2;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 CC Rhabditidae; Pelodierinae; Caenorhabditis.
 CC NCBI_Taxid=6239;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Bristol N2;
 RX MEDLINE=9069613; PubMed=9851916;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology."
 RT Science 283:2012-2018 (1998).

DR EMBL; Z49068; CAA8885.1; -; Genomic_DNA.
 DR PIR; T23167; T23167.
 DR Ensembl; K01C8.2; Caenorhabditis elegans.
 DR Wormbase; WBGene00010457; K01C8.2.
 DR Wormpep; K01C8.2; C802263.
 DR InterPro; IPR006150; Worm_repeat_1.
 DR SMART; SM00289; WR1; 7.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 389 AA; 40335 MW; C32B38446991D17D CRC64;

Query Match 65.4%; Score 17; DB 2; Length 389;
 Best Local Similarity 20.0%; Pred. No. 2.8;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXC 10
 DB 123 CSSSISTSSC 132

RESULT 31

Q67U23_ORYZA PRELIMINARY; PRT; 411 AA.

AC Q67U23; 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Chloroplast thylakoidal processing peptidase-like protein.
 GN Name=P0488D02.23;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Ehrhartoideae; Oryzaceae; Oryza.
 CC NCBI_Taxid=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC
 RT clone: P0488D02."
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005393; BAD38026.1; -; Genomic_DNA.
 DR Gramene; Q67U23;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR CO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000223; Peptidase_S26A.
 DR InterPro; IPR006198; Peptidase_S26.
 DR Pfam; PF00717; Peptidase_S24; 1.
 DR PRINTS; PR00727; LEADERPTASE.
 DR TIGRfam; TIGR02227; sigsep_1_bact; 1.
 DR PROSITE; PS00501; SPASE_1; 1.
 DR PROSITE; PS00761; SPASE_1_3; 1.
 KW Hydrolase.

SQ SEQUENCE 411 AA; 43573 MW; EE76AD564816A373 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 411;
 Best Local Similarity 20.0%; Pred. No. 2.9;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXC 10
 DB 25 CSSASTAC 34

RESULT 32

Q41408_GIBZE PRELIMINARY; PRT; 437 AA.

AC Q41408; 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=FG07600.1;
 OS Gibberella zeae PH-1.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 CC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
 CC NCBI_Taxid=229533;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PH-1;
 RA Birren B., Nuebaum C., Abouelleil A., Allen N., Anderson S.,
 RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavsky L.,
 RA Boukhalter B., Butler J., Calvo S.E., Camarata J., Chang J.,

RA Choepel Y., Collamore A., Cook A., Cooke P., Corum B., Dearrellano K.,
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
 RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
 RA Gardyna S., Gierre S., Graham L., Grand-Pierre N., Hafez N.,
 RA Hagopian D., Hages B., Hall J., Horton L., Hulme W., Iliev I.,
 RA Jaffe D., Johnson R., Jones C., Kamat A., Karatas A.,
 RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
 RA Ma L.-Y., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
 RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Meneus L.,
 RA Milnova T., Menga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
 RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
 RA Oliver J., Peterson K., Phunkhang P., Pierre N., Percell S.,
 RA Rachuapa A., Ramsamy U., Raymond C., Retta R., Rise C., Rogov P.,
 RA Roman J., Schauer S., Schupack R., Seaman S., Severy P., Smirnov S.,
 RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs W.,
 RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
 RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
 RA Lander E.,
 RT "Fusarium graminearum genome sequence."
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL, AACM01000320; EAA77994.1; -; Genomic_DNA.
 KM Hypothetical protein.
 SQ SEQUENCE 437 AA; 47817 MW; 01EBA0260503394 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 437;
 Best Local Similarity 20.0%; Pred. No. 2.9;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 10
 Db 199 CATLASTSSC 208

RESULT 33
 Q9VMG7 DROME PRELIMINARY; PRT; 471 AA.
 ID Q9VMG7 DROME PRELIMINARY; PRT; 471 AA.
 AC Q9VMG7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE CG13990-PA.
 OS Name=CG13990; ORFNames=CG13990;
 GN Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydriidae; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Balles R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew J., Dietz S.M.,
 RA Doudon K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gelfand A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke J., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Liao P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulyov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reiter K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu K., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195 (2000).

[2]
 RN NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celinker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Taber S.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence."
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

[3]
 RN NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celinker S.B.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomic perspective."
 RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).

[4]
 RN NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Celinker S.B., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review."
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).

[5]
 RN NUCLEOTIDE SEQUENCE.
 RG Berkeley Drosophila Genome Project;
 RA Celinker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
 RA Hoskins R., Stapleton M., Pacle J., Park S., Svirskas R., Smith E.,
 RA Yu C., Rubin G.;
 RT "Drosophila melanogaster release 4 sequence."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[6]
 RN NUCLEOTIDE SEQUENCE.
 RP Flybase;
 RG Flybase;
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AEO03612; AAF52351.2; -; Genomic DNA.
 DR Ensembl; CG13990; Drosophila melanogaster.
 DR Flybase; FBgn0040950; CG13990.
 DR GO; GO:0005576; C:extracellular region; IEA.
 DR GO; GO:0008061; P:chitin binding; IEA.
 DR GO; GO:0006030; P:chitin metabolism; IEA.
 DR InterPro; IPR002557; Chitin_bind_Pera.
 DR Pfam; PF01607; CBM_14; 2.

DR SMART; SM00494; ChtBD2; 2.
DR PROSITE; PSS0940; CHIT_BIND II; 2.
SQ SEQUENCE 471 AA; 4811 MW; F3167EB0728B9578 CRC64;
Query Match 65.4%; Score 17; DB 2; Length 471;
Best Local Similarity 20.0%; Pred. No. 3;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 1 CXXXXXXXXX 10
DB 350 CASSSTTTC 359
RESULT 34
O7X7A4 ORYSA PRELIMINARY; PRT; 475 AA.
ID O7X7A4_ORYSA PRELIMINARY; PRT; 475 AA.
AC O7X7A4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE OSJNB0060P14.14 protein (OSJNB0048E02.5 protein).
CN Name=OSJNB0060P14.14; Synonyms=OSJNB0048E02.5;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Yu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Lan L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL663017; CAE04361.2; -; Genomic DNA.
DR EMBL; AL606653; CAE04825.2; -; Genomic DNA.
DR Gramene; O7X7A4; -;
DR GO; GO:0006512; P;ubiquitin cycle; IEA.
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00256; FBOX; 1.
DR PROSITE; PSS0181; FBOX; 1.
SQ SEQUENCE 475 AA; 53063 MW; 63345B3CB2D4CA5 CRC64;
Query Match 65.4%; Score 17; DB 2; Length 475;
Best Local Similarity 20.0%; Pred. No. 3;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 1 CXXXXXXXXX 10
DB 417 CASSSTLSSAC 426
RESULT 35
O4P4C7 USTMA PRELIMINARY; PRT; 491 AA.
ID O4P4C7_USTMA PRELIMINARY; PRT; 491 AA.
AC O4P4C7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM05036.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
[1]

OX NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Alt-zania M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantang P., Baldwin J., Barry A.,
RA Bayati T., Blitshteyn B., Bloom T., Blye J., Boguslavsky L.,
RA Borowsky M., Boukhalter B., Brunache A., Butler J., Caliste N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshtang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Dufley N., Dupes A., Ekins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gritke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hages B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Huey E., Iliev I.,
RA Jaffe D., Jones C., Kamel M., Kamet A., Kamyssele M., Karlsson E.,
RA Kells C., Kieu A., Kiser P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokytesang T., Lokytesang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabelli R., Maru K., Matthews C., Mauceli E.,
RA McCarthy W., McDonough S., Mcghee T., Meldrum J., Meneus L.,
RA Mesirov J., Mihalov A., Minova T., Mikkelsen T., Mianga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., Neves C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotombo B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piganì B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Rectori S., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schubach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Souarez C.,
RA Spencer B., Stalker J., Strange-thomann N., Stavrakopoulos S.,
RA Stenson P., Stone C., Stone S., Stubbs M., Talamas Y., Tchuinga P.,
RA Teasing P., Testaye S., Theodore J., Thoulustang Y., Topham K.,
RA Towey S., Tsamila T., Tsomo N., Vallée D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Vengkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whitlaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zairoun J., Zembeck L.,
RA Zimmer A., Zody M., Zander B.;
RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACP0100186; EAK85896.1; -; Genomic DNA.
DR Hypothetical protein.
SQ SEQUENCE 491 AA; 51869 MW; 768A84FB8BD71BBE CRC64;
Query Match 65.4%; Score 17; DB 2; Length 491;
Best Local Similarity 20.0%; Pred. No. 3.1;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 1 CXXXXXXXXX 10
DB 453 CASSSTGSSSC 462
RESULT 36
O4XZV5 PLACH PRELIMINARY; PRT; 520 AA.
ID O4XZV5_PLACH PRELIMINARY; PRT; 520 AA.
AC O4XZV5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
GN ORFNames=PC000338.02.0;
OS Plasmodium chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5825;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karris M., Raine J.D., Carlton J.M., Kooij T.M.A.,
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutledge K., Harris B., Harris D., Church C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carnucci D.J., Yates J.R., Kafatos F.C.,
RA Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.,
RT "A comprehensive survey of the plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses."
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: CAJ01002036; CAJ77555.1; -; Genomic_DNA.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 5.
DR PRINTS: PR00320; GPROTEINREPT.
DR SMART: SM00320; WD40; 5.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_1.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
DR Hypothetical protein; Repeat; WD repeat.
KW NON_TER 520 520
FT SEQUENCE 520 AA; 59951 MW; FA789C11BFD99087 CRC64;
SQ
Query Match 65.4%; Score 17; DB 2; Length 520;
Best Local Similarity 20.0%; Pred. No. 3.1;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
DB 364 CSTSLSSSC 373

RESULT 37
O803D5 BRARE PRELIMINARY; PRT; 556 AA.
AC O803D5
ID O803D5
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to adrenal secretory serine protease.
GN ORFNames=zsc:55888;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN-AB; TISSUE=whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schlier G.D.,
RA Alechul S.F., Zeeberg B., Buettow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stabileton M., Soares M.B., Bonaldo W.F., Caesavanti T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rahn S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettlemen M., Madan A., Rodighiero S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalek U., Smalls D.E.,
RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN-AB; TISSUE=whole body;
RA Strausberg R.,
RT Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC044526; AAH44526.1; -; mRNA.
DR HSSP: P00760; IEZK.
DR Ensembl: ENSDARG0000016538; Danio rerio.
DR ZFIN: ZDB-GENE-040426-834; zgc:55888.
DR GO: GO:0004263; F:chymotrypsin activity; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0004295; F:trypsin activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR001254; Peptidase SL_S6.
DR InterPro: IPR001314; Peptidase SLA.
DR Pfam: PF00089; Trypsin, 2.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRYPSIN_DOM; 2.
DR PROSITE: PS50240; TRYPSIN_DOM; 2.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
KW SEQUENCE 556 AA; 61077 MW; C7B23D930547878C CRC64;
SQ
Query Match 65.4%; Score 17; DB 2; Length 556;
Best Local Similarity 20.0%; Pred. No. 3.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
DB 493 CTHAAASTSC 502

RESULT 38
O5GXM1_XANOR
ID O5GXM1_XANOR PRELIMINARY; PRT; 568 AA.
AC O5GXM1;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Alkaline phosphatase.
GN Name=phoA; OrderedLocustNames=XO03296;
OS Xanthomonas oryzae (pv. oryzae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=64187;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=RACC10331 / KX085;
RX PubMed=15673718; DOI=10.1093/nar/gk1206;
RA Lee B.-M., Park Y.-U., Park D.-S., Kang H.-W., Kim J.-G., Song E.-S.,
RA Park I.-C., Yoon U.-H., Hahn J.-H., Ko B.-S., Lee G.-B., Kim H.,
RA Park H.-S., Yoon K.-O., Kim J.-H., Jung C.-H., Koh N.-H., Seo J.-S.,
RA Go S.-U.;
RT "The genome sequence of Xanthomonas oryzae pathovar oryzae RACC10331,
RT the bacterial blight pathogen of rice."
RL Nucleic Acids Res. 33:577-586(2005).
DR EMBL: AE013598; AAW7550.1; -; Genomic_DNA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR001952; ALK_phosphatase.
DR Pfam: PF00245; ALK_phosphatase; 1.
DR PRINTS: PR00113; ALKPHPTASE.
DR SMART: SM00098; alkppc; 1.
KW Complete proteome.
KW SEQUENCE 568 AA; 60106 MW; CA69E266ACB3A6 CRC64;
SQ
Query Match 65.4%; Score 17; DB 2; Length 568;
Best Local Similarity 20.0%; Pred. No. 3.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
DB 21 CASTAGSSAC 30

```

RESULT 39
ID Q4N2Y7 THEPA PRELIMINARY; PRT; 626 AA.
AC Q4N2Y7;
DT 13-SEP-2005 (TEMBLrel. 31, Created)
DT 13-SEP-2005 (TEMBLrel. 31, Last sequence update)
DE Hypothetical protein.
OS ORFNames=TP04_0205;
GN Theileria parva.
OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileridae;
OC Theileria.
NCBI_Taxid=5875;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Muga;
RA Gardner M.V., Bishop R., Shah T., de Villiers E.P., Carlton J.M.,
RA Hall N., Ren Q., Paulsen I.T., Pain A., Berriman M., Wilson R.J.M.,
RA Sato S., Ralph S.A., Mann D.J., Xiong Z., Shalim S.J., Weidman J.,
RA Jiang L., Lynn J., Weaver B., Shoabli A., Masawo D., Crabtree J., Mortman J.R.,
RA Mortman J.R., Haas B., Angiuoli S.V., Creasy T.H., Lu C., Sun B.,
RA Silva J.C., Uterback T.R., Feldblyum T.V., Perce M., Allen J.,
RA Taracha E.L.N., Salzberg S.L., White O.R., Fitzhugh H.A., Morzaria S.,
RA Venter J.C., Fraser C.M., Nene V.;
RT "Genome sequence of Theileria parva, a bovine pathogen that transforms
RT lymphocytes."
RL Science 309:134-137(2005).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Muga;
RA Gardner M., Bishop R., Shah T., de Villiers E., Carlton J.M., Hall N.,
RA Ren Q., Paulsen I.T., Pain A., Berriman M., Wilson R.J.M., Sato S.,
RA Ralph S.A., Mann D.J., Xiong Z., Shalim S.J., Weidman J., Jiang L.,
RA Lynn J., Weaver B., Shoabli A., Masawo D., Crabtree J., Mortman J.R.,
RA Haas B., Angiuoli S., Creasy T.H., Lu C., Sun B., Silva J.C.,
RA Uterback T., Feldblyum T., Perce M., Allen J., Taracha E.L.,
RA Salzberg S.L., White O., Fitzhugh H.A., Morzaria S., Venter J.C.,
RA Fraser C.M., Nene V.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC DR EMBL; AAGK0100004; EAX31557.1; -; Genomic DNA.
KW Hypothetical protein.
SQ
SEQUENCE 626 AA; 72615 MW; 663085BA95D09916 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 626;
Best Local Similarity 20.0%; Pred. No. 3.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 10
DB 44 CAASSSDASC 53

RESULT 40
ID Q69TY8 ORYSA PRELIMINARY; PRT; 630 AA.
AC Q69TY8;
DT 25-OCT-2004 (TEMBLrel. 28, Created)
DT 25-OCT-2004 (TEMBLrel. 28, Last sequence update)
DE 10-MAY-2005 (TEMBLrel. 30, Last annotation update)
DE Receptor protein kinase-like.
GN Name=OSUNBa0029G06.34; Synonyms=OSUNBa0072A21.3;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Euphorbiaceae; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Euphorbiaceae; Oryzae; Oryza.
OX NCBI_Taxid=39947;
RN
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC

```

```

RT clone:OSUNBa0029G06.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC
RL clone:OSUNBa0072A21.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004680; BAD35689.1; -; Genomic DNA.
DR EMBL; AP004737; BAD37734.1; -; Genomic DNA.
DR Gramene; Q69TY8;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR002482; LysM.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF01476; LysM; 1.
DR Pfam; PF00069; Kinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
KW Kinase; Receptor.
SQ
SEQUENCE 630 AA; 67017 MW; F0EFCB8BBA63646 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 630;
Best Local Similarity 20.0%; Pred. No. 3.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 10
DB 48 CSEYATATSC 57

RESULT 41
ID HSP71 LEIMA STANDARD; PRT; 634 AA.
AC P12076;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-FEB-2005 (Rel. 46, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Heat shock 70-related protein 1, mitochondrial precursor.
GN Name=HSP70.1;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_Taxid=5664;
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MHOM/IL/81/Friedlin;
RA MEDLINE=93300981; PubMed=8314893;
RA Searle S., McCrossan M.V., Smith D.F.;
RT "Expression of a mitochondrial stress protein in the protozoan
RT parasite Leishmania major.";
RL U. Cell Sci. 104:1091-1100(1993).
[2]
RP SEQUENCE REVISION TO 461-500.
RA Smith D.F.;
RN Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
[3]
RP NUCLEOTIDE SEQUENCE OF 1-249.
RC STRAIN=MHOM/IL/81/Friedlin;
RA MEDLINE=89345072; PubMed=2762121;
RA Searle S., Campos A.J.R., Coulson R.M.R., Spithill T.W., Smith D.F.;
RT "A family of heat shock protein 70-related genes are expressed in the
RT promastigotes of Leishmania major.";
RL Nucleic Acids Res. 17:5081-5095(1989).
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.

```

```
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X64137; CAA5498.2; -; Genomic DNA.
DR EMBL; X14574; CAA32713.1; -; Genomic DNA.
DR PIR; S33575; S33575.
DR HSSP; P04475; 1DKG.
DR InterPro; IPR012725; Dnak prot.
DR InterPro; IPR0101023; Hsp70.
DR PANTHER; PTHR19375; Hsp70; 1.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR ProDom; PD000089; Hsp70; 1.
DR TIGRFAMs; TIGR02350; prok_dnak; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; FALSE_NEG.
DR ATP-binding; Coiled coil; Heat shock; Mitochondrion; Multigene family;
DR Nucleotide-binding; Transist peptide.
FT TRANSIT 1 634 Mitochondrion (By similarity).
FT CHAIN 1 634 Heat shock 70-related protein 1.
FT COLLED 538 614 Potential.
SQ SEQUENCE 634 AA; 68330 MW; 8382B1C40C108EEA CRC64;

Query Match
Best Local Similarity 65.4%; Score 17; DB 1; Length 634;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
Db 7 CGSAAASAAAC 16

RESULT 42
Q0747_LEIMA
ID Q0747_LEIMA PRELIMINARY; PRT; 635 AA.
AC Q0747;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Heat shock 70-related protein 1, mitochondrial, putative.
GN ORFNames=LmjF30.2460;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
DR EMBL; CT005267; CAJ06531.1; -; Genomic DNA.
DR InterPro; IPR012725; Dnak prot.
DR InterPro; IPR0101023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR ProDom; PD000089; Hsp70; 1.
DR TIGRFAMs; TIGR02350; prok_dnak; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; 1.
DR ATP-binding; Heat shock; Nucleotide-binding.
KW SEQUENCE 635 AA; 68948 MW; 1A11F69117A2200 CRC64;

Query Match
Best Local Similarity 65.4%; Score 17; DB 2; Length 635;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
Qy 1 CXXXXXXXXXC 10
Db 7 CGSAAASAAAC 16

RESULT 43
Q0740_LEIMA
ID Q0740_LEIMA PRELIMINARY; PRT; 652 AA.
AC Q0740;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Heat shock 70-related protein 1, mitochondrial, putative.
GN ORFNames=LmjF30.2550;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
DR EMBL; CT005267; CAJ06531.1; -; Genomic DNA.
DR InterPro; IPR012725; Dnak prot.
DR InterPro; IPR0101023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.

Query Match
Best Local Similarity 65.4%; Score 17; DB 2; Length 652;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10
Db 7 CGSAAASAAAC 16

RESULT 44
Q0744_LEIMA
ID Q0744_LEIMA PRELIMINARY; PRT; 660 AA.
AC Q0744;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Heat shock 70-related protein 1, mitochondrial, putative.
GN ORFNames=LmjF30.2490;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
DR EMBL; CT005267; CAJ06531.1; -; Genomic DNA.
DR InterPro; IPR012725; Dnak prot.
DR InterPro; IPR0101023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
```


DR Prodom; PD000089; Hsp70; 1.
 DR TIGRPMs; TIGR02350; prok_dnak; 1.
 DR PROSITE; PS00297; HSP70_1; 1.
 DR PROSITE; PS00329; HSP70_2; 1.
 DR PROSITE; PS01036; HSP70_3; 1.
 KW ATP-binding; Heat shock; Nucleotide-binding.
 SQ SEQUENCE 660 AA; 71621 MW; 5E1D306B3D524242 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 660;
 Best Local Similarity 20.0%; Pred. No. 3.4;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10
 Db 7 CGSAAASAC 16

RESULT 45
 ID Q40745 LEIMA PRELIMINARY; PRT; 662 AA.

AC Q40745_1
 DT 13-SEP-2005 (TREMBlrel. 31, Created)
 DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
 DE Heat shock 70-related protein 1, mitochondrial, putative.
 GN ORFNames=LmjF30.2470, LmjF30.2480;
 OS Leishmania major.
 OC Eukaryota; Buglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxId=5664;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Friedlin;
 RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
 RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
 RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
 DR EMBL; C005267; CAJ06535.1; -; Genomic DNA.
 DR EMBL; C005267; CAJ06535.1; -; Genomic DNA.
 DR InterPro; IPR012725; Dnak_prot.
 DR InterPro; IPR001023; Hsp70.
 DR Pfam; PF00012; HSP70; 1.
 DR PROSITE; PRO0301; HEATSHOCK70.
 DR Prodom; PD000089; Hsp70; 1.
 DR TIGRPMs; TIGR02350; prok_dnak; 1.
 DR PROSITE; PS00297; HSP70_1; 1.
 DR PROSITE; PS00329; HSP70_2; 1.
 DR PROSITE; PS01036; HSP70_3; 1.
 KW ATP-binding; Heat shock; Nucleotide-binding.
 SQ SEQUENCE 662 AA; 71877 MW; CF3C66350EC46E20 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 662;
 Best Local Similarity 20.0%; Pred. No. 3.4;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10
 Db 7 CGSAAASAC 16

RESULT 46

ID Q60E18 ORYSA PRELIMINARY; PRT; 691 AA.
 AC Q60E18_1
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DE Putative cyclic nucleotide-gated ion channel.
 GN Name=OSUNBA0017K09.10;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.

OX NCBI_TaxId=39947;

RP NUCLEOTIDE SEQUENCE.
 RA Chow T.-Y., Heing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
 RA Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
 RA Chen Y.-L., Cheng C.-H., Chung C.-H., Han S.-Y., Hsiao S.-H.,
 RA Heitung J.-N., Hsu C.-H., Huang J.-Y., Kau P.-I., Lee M.-C., Leu H.-L.,
 RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
 RA Wu H.-P., Shaw J.-F.;
 RL "Oryza sativa BAC OSUNBA0017K09 genomic sequence";
 RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC130597; AAU90233.1; -; Genomic DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005216; P:ion channel activity; IEA.
 DR GO; GO:0006811; P:ion transport; IEA.
 DR InterPro; IPR005953; cNMP_binding.
 DR InterPro; IPR005821; Ion_trans.
 DR Pfam; PF00027; cNMP_binding; 1.
 DR Pfam; PF00520; Ion_trans; 1.
 DR SMART; SM00100; cNMP; 1.
 DR PROSITE; PS50042; cNMP_BINDING_3; 1.
 SQ SEQUENCE 691 AA; 77220 MW; D059333316F2E299 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 691;
 Best Local Similarity 20.0%; Pred. No. 3.5;
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10
 Db 298 CASSTAANVAC 307

RESULT 47

ID Q819P8 9BIYA PRELIMINARY; PRT; 699 AA.
 AC Q819P8_1
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DE "Horizontal gene transfer from Burkarya to bacteria and domain shuffling: the alpha-amyase model."
 DE Alpha-amyase.
 GN Name=amy;
 OS Corbicula fluminea.
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
 OC Corbiculoidea; Corbiculidae; Corbicula.
 OX NCBI_TaxId=45949;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RA Da Lage J.-L., Van Wormhoudt A., Carlot M.-L.;
 RT "Diversity and evolution of the alpha-amyase genes in Animals."
 RL Biologia 57:181-189(2002).
 RN (2)

RP NUCLEOTIDE SEQUENCE.
 RX PubMed=14704857; DOI=10.1007/s00018-003-3334-Y;
 RA Da Lage J.-L., Feller G., Janacek S.;
 RT "Horizontal gene transfer from Burkarya to bacteria and domain shuffling: the alpha-amyase model."
 RL Cell. Mol. Life Sci. 61:97-109(2004).
 RN (3)
 RP NUCLEOTIDE SEQUENCE.
 RA Da Lage J.-L.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF468016; ANO17927.2; -; Genomic DNA.
 DR HSSP; P04745; 1SMD.
 DR GO; GO:0004556; P:alpha-amyase activity; IEA.
 DR GO; GO:0005509; P:calcium ion binding; IEA.
 DR GO; GO:0016798; P:hydrolase activity, acting on glycosyl bonds; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR006048; Alpha_amy1_C.
 DR InterPro; IPR006047; Alpha_amy1_cat.
 DR InterPro; IPR006589; Alp_amy1_cat_sub.
 DR InterPro; IPR006046; Glyco_hydro_13.
 DR Pfam; PF00128; Alpha-amyase; 1.
 DR Pfam; PF02806; Alpha-amyase_C; 1.

```

DR PRINTS: PR00110; ALPHAAMYLASE.
DR SMART; SM00642; Aamy; 1.
DR SMART; SM00632; Aamy; C; 1.
SQ SEQUENCE 699 AA; 76544 MW; 25D57008165B04CB CRC64;

Query Match
Best Local Similarity 20.0%; Score 17; DB 2; Length 699;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
DB 544 CTSDATSAC 553

RESULT 48
O5L6X2_CHLAB
ID O5L6X2_CHLAB PRELIMINARY; PRT; 705 AA.
AC O5L6X2;
DT 01-FEB-2005 (TReMBLrel. 29, Created)
DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)
DE Putative transport protein.
GN OrderedLocusNames=CAB140;
OS Chlamydomophila abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=83555;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC PubMed=15837807; DOI=10.1101/gr.3684805;
RA Thomson N.R., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,
RA Livingstone M., Cerdano-Tarraga A.-M., Harris B., Doggett J.,
RA Ormond D., Mungall K., Clarke K., Felwell T., Hance Z., Sanders M.,
RA Quail M.A., Price C., Barrett B.G., Parkhill J., Longbottom D.;
RT "The Chlamydomophila abortus genome sequence reveals an array of
RT variable proteins that contribute to interspecies variation.";
RL Genome Res. 15:629-640 (2005).
DR EMBL; CR848038; CAH63598.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 705 AA; 81369 MW; 1997374B5266E2AB CRC64;

Query Match
Best Local Similarity 20.0%; Score 17; DB 2; Length 705;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
DB 67 CSSLSATTC 76

RESULT 49
O9NKE3_DROME
ID O9NKE3_DROME PRELIMINARY; PRT; 734 AA.
AC O9NKE3;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE Hypothetical protein BG:DS00180.5.
GN Name=BG:DS00180.5; ORFNames=CG31765, CG31766, CG31841;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=99403001; PubMed=10471707;
RA Ashburner M., Mista S., Roote J., Lewis S.E., Blazee R.G., Davis T.,
RA Doyle C., Galle R.F., George R.A., Harris N.L., Hartzell G.,
RA Harvey D.A., Hong L., Houston K.A., Hoskins R.A., Johnson G.,
RA Martin C., Moshrefi A.R., Palazzolo M., Reese M.G., Spradling A.C.,
RA Teang G., Wan K.H., Whitelaw K., Celniker S.E., Rubin G.M.;

```

```

RT "An exploration of the sequence of a 2.9-Mb region of the genome of
RT Drosophila melanogaster: the adh region.";
RL Genetics 153:179-219(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkeley;
RA Celniker S.E., Agbayani A., Arcaina T.T., Baxter B., Blazee R.G.,
RA Butenhof C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Fartan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummachi S.R., Karia K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomolan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA Sethi H., Snit E., Swirekas R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003408; AAF44838.1; -; Genomic_DNA.
DR FlyBase; FBgn0051765; CG31765.
DR FlyBase; FBgn0051766; CG31766.
DR FlyBase; FBgn0051841; CG31841.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004245; P:peptidyl-dipeptidase A activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR01548; Peptidase_M2.
DR Pfam; PF01401; Peptidase_M2; 1.
KW Hypothetical protein.
SQ SEQUENCE 734 AA; 79831 MW; 4C0B5993AF9PBC1E CRC64;

Query Match
Best Local Similarity 20.0%; Score 17; DB 2; Length 734;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
DB 681 CSSASAVASC 690

RESULT 50
O4YNR2_PLABE
ID O4YNR2_PLABE PRELIMINARY; PRT; 784 AA.
AC O4YNR2;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PB000881.03.0;
OS Plasmodium berghei.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5821;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karras M., Raine J.D., Carlton J.M., Koof T.W.A.,
RA Bertman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churchar C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Jans C.J., Barrett B., Turner C.M.R., Waters A.P., Sindern R.S.;
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses.";
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CA01003291; CA100348.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 784 AA; 91429 MW; B76DA6BAD7A0C1D6 CRC64;

Query Match
Best Local Similarity 20.0%; Score 17; DB 2; Length 784;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
DB 363 CSSSFSFSSC 372

```

Thu Jan 5 09:14:13 2006

us-09-932-322-11.rup

Page 25

Search completed: January 4, 2006, 16:09:55
Job time : 67.3043 secs

4444 01234567

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 : Search time 82.1217 Seconds
(without alignments)
64.204 Million cell updates/sec

Title: US-09-932-322-12
Perfect score: 28
Sequence: 1 CXXXXXXXXXXC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq.21.*
1: geneseqp1980s:*
2: geneseqp2000s:*
3: geneseqp2001s:*
4: geneseqp2002s:*
5: geneseqp2003s:*
6: geneseqp2004s:*
7: geneseqp2005s:*
8: geneseqp2006s:*
9: geneseqp2007s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	64.3	104	7	ABO82562 Pseudomon
2	18	64.3	110	8	ADP30723 Human sec
3	18	64.3	135	8	ADP31638 Human sec
4	18	64.3	142	7	ABO79745 Pseudomon
5	18	64.3	178	7	ABO73284 Pseudomon
6	18	64.3	264	8	ADP31527 Human sec
7	18	64.3	270	8	ADP31435 Human sec
8	18	64.3	307	4	ABO65879 Drosophil
9	18	64.3	357	8	ADP31223 Human sec
10	18	64.3	615	8	ADP31132 Human sec
11	18	64.3	627	6	ABM48099 Propionib
12	18	64.3	627	6	ABM48099 Propionib
13	18	64.3	669	8	ADP31142 Human sec
14	18	64.3	711	8	ADP31215 Human sec
15	18	64.3	739	8	ADP31196 Human sec
16	18	64.3	771	8	ADP31244 Human sec
17	18	64.3	876	8	ADP31220 Human sec
18	18	64.3	882	8	ADP31688 Human sec
19	18	64.3	925	5	AAO14246 Human pre
20	18	64.3	1044	8	ADP31517 Human sec
21	18	64.3	1086	8	ADP31175 Human sec
22	18	64.3	1113	8	ADP31508 Human sec
23	18	64.3	1128	6	ADA15725 C. elegan
24	18	64.3	1134	8	ADP30741 Human sec

25	18	64.3	1134	8	ADP30924 Human sec
26	18	64.3	1168	8	ADP31046 Human sec
27	18	64.3	1191	8	ADP30993 Human sec
28	18	64.3	1260	8	ADP31533 Human sec
29	18	64.3	1289	8	ADP30675 Human sec
30	18	64.3	1337	8	ADP31357 Human sec
31	18	64.3	1454	8	ADP31177 Human sec
32	18	64.3	1480	8	ADP30557 Human sec
33	18	64.3	1588	5	ABO69437 H. influenza
34	18	64.3	1617	8	ADP30660 Human sec
35	18	64.3	1652	6	ADA15715 C. elegan
36	18	64.3	1725	8	ADP30654 Human sec
37	18	64.3	1833	8	ADP30642 Human sec
38	18	64.3	2088	8	ADP31178 Human sec
39	18	64.3	2127	8	ADP31327 Human sec
40	18	64.3	2484	8	ADP66690 Human mis
41	18	64.3	2508	8	ADA15721 C. elegan
42	18	64.3	2544	6	ADA15717 C. elegan
43	18	64.3	2601	6	ADA15723 C. elegan
44	18	64.3	2833	8	ADP31299 Human sec
45	18	64.3	2835	8	ADP30572 Human sec
46	18	64.3	3411	8	ADP30667 Human sec
47	18	64.3	4440	6	ABO88256 Human sec
48	18	64.3	4440	6	ABO90135 Human sec
49	18	64.3	4440	6	ABO96437 Human sec
50	18	64.3	4440	6	ABO99046 Human sec
51	18	64.3	4440	6	ABO98261 Human sec
52	18	64.3	4440	6	ABO91967 Human sec
53	18	64.3	4440	6	ABO85271 Human sec
54	18	64.3	4440	6	ABO00410 Human sec
55	18	64.3	4440	6	ABO88961 Human sec
56	18	64.3	4440	6	ABO06457 Human sec
57	18	64.3	4440	6	ABO95517 Human sec
58	18	64.3	4440	6	ABO95207 Human sec
59	18	64.3	4440	6	ABO90755 Human sec
60	18	64.3	4440	6	ABO93917 Human sec
61	18	64.3	4440	6	ABO86191 Human sec
62	18	64.3	4440	6	ABO82046 Human sec
63	18	64.3	4440	6	ABO07907 Human sec
64	18	64.3	4440	6	ABO94227 Human sec
65	18	64.3	4440	6	ABO00100 Human sec
66	18	64.3	4440	6	ABO87111 Human sec
67	18	64.3	4440	6	ABO91352 Human sec
68	18	64.3	4440	6	ABO90445 Human sec
69	18	64.3	4440	6	ABO97036 Human sec
70	18	64.3	4440	6	ABO05232 Human sec
71	18	64.3	8976	8	ADP31425 Human sec
72	18	64.3	9195	8	ADP31494 Human sec
73	18	64.3	9195	8	ADP30534 Human sec
74	18	60.7	31	8	ADP30526 Human sec
75	18	60.7	53	3	AGS58907 Arabidops
76	18	60.7	65	4	ABG29760 Arabidops
77	18	60.7	95	8	ADP30709 Human sec
78	18	60.7	105	7	ABO78139 Human sec
79	18	60.7	119	7	ABO79679 Human sec
80	18	60.7	149	3	AA94985 Human sec
81	18	60.7	160	8	ADP31048 Human sec
82	18	60.7	165	8	ADP31515 Human sec
83	18	60.7	165	8	ADP31516 Human sec
84	18	60.7	168	7	ABO79318 Human sec
85	18	60.7	168	8	ADP31621 Human sec
86	18	60.7	168	8	ADP31099 Human sec
87	18	60.7	181	8	ADP32231 Human sec
88	18	60.7	234	7	ABO68370 Human sec
89	18	60.7	240	8	ADP30737 Human sec
90	18	60.7	252	8	ADP31487 Human sec
91	18	60.7	264	7	ABO75516 Human sec
92	18	60.7	264	7	ADP31272 Human sec
93	18	60.7	297	8	ADP31312 Human sec
94	18	60.7	309	8	ADP30862 Human sec
95	18	60.7	309	8	ADP30873 Human sec
96	18	60.7	320	8	ADP31607 Human sec
97	18	60.7	320	8	ADP31649 Human sec

98	17	60.7	325	7	ABM87858	Abm87858	Rice	ab10	171	17	60.7	950	8	ADP31167	Adp31167	Human	sec
99	17	60.7	330	7	ADB70062	Adb70062	C. neofor		172	17	60.7	951	8	ADP30943	Adp30943	Human	sec
100	17	60.7	348	8	ADP31441	Adp31441	Human	sec	173	17	60.7	957	8	ADP31528	Adp31528	Human	sec
101	17	60.7	357	8	ADP31267	Adp31267	Human	sec	174	17	60.7	957	8	ADP07508	Adp07508	Human	col
102	17	60.7	367	8	ADP30950	Adp30950	Human	sec	175	17	60.7	960	8	ADP31471	Adp31471	Human	sec
103	17	60.7	381	8	ADP30655	Adp30655	Human	sec	176	17	60.7	960	8	ADP31470	Adp31470	Human	sec
104	17	60.7	384	8	ADP30656	Adp30656	Human	sec	177	17	60.7	981	8	ADP30547	Adp30547	Human	sec
105	17	60.7	392	8	ADP31061	Adp31061	Human	sec	178	17	60.7	996	8	ADP31538	Adp31538	Human	sec
106	17	60.7	392	8	ADP31055	Adp31055	Human	sec	179	17	60.7	1035	8	ADP31552	Adp31552	Human	sec
107	17	60.7	393	8	ADP31345	Adp31345	Human	sec	180	17	60.7	1038	8	ADP30860	Adp30860	Human	sec
108	17	60.7	402	8	ADP30742	Adp30742	Human	sec	181	17	60.7	1044	8	ADP31550	Adp31550	Human	sec
109	17	60.7	402	8	ADP31414	Adp31414	Human	sec	182	17	60.7	1048	8	ADP31642	Adp31642	Human	sec
110	17	60.7	402	8	ADP31262	Adp31262	Human	sec	183	17	60.7	1053	8	ADP30886	Adp30886	Human	sec
111	17	60.7	409	7	ABO79810	AbO79810	Pseudomon		184	17	60.7	1059	8	ADP31042	Adp31042	Human	sec
112	17	60.7	411	8	ADP30729	Adp30729	Human	sec	185	17	60.7	1065	8	ADP31347	Adp31347	Human	sec
113	17	60.7	412	8	ADP57242	AdP57242	Plant	pol	186	17	60.7	1076	6	ABU81145	Abu81145	Human	PRO
114	17	60.7	420	8	ADP31273	Adp31273	Human	sec	187	17	60.7	1076	6	ABU66845	Abu66845	Human	PRO
115	17	60.7	420	8	ADP31274	Adp31274	Human	sec	188	17	60.7	1082	8	ADP30934	Adp30934	Human	sec
116	17	60.7	423	8	ADP31479	Adp31479	Human	sec	189	17	60.7	1092	8	ADP31358	Adp31358	Human	sec
117	17	60.7	483	8	ADP30852	Adp30852	Human	sec	190	17	60.7	1101	8	ADP31462	Adp31462	Human	sec
118	17	60.7	483	8	ADP30853	Adp30853	Human	sec	191	17	60.7	1110	8	ADP31430	Adp31430	Human	sec
119	17	60.7	510	8	ADP31049	Adp31049	Human	sec	192	17	60.7	1119	8	ADP31452	Adp31452	Human	sec
120	17	60.7	525	8	ADP31227	Adp31227	Human	sec	193	17	60.7	1128	8	ADP30483	Adp30483	Human	sec
121	17	60.7	528	8	ADP30529	Adp30529	Human	sec	194	17	60.7	1134	8	ADP30647	Adp30647	Human	sec
122	17	60.7	531	8	ADP31539	Adp31539	Human	sec	195	17	60.7	1134	8	ADP31537	Adp31537	Human	sec
123	17	60.7	531	8	ADP31540	Adp31540	Human	sec	196	17	60.7	1141	8	ADP30677	Adp30677	Human	sec
124	17	60.7	543	8	ADP30864	Adp30864	Human	sec	197	17	60.7	1170	8	ADP30819	Adp30819	Human	sec
125	17	60.7	552	8	ADP31019	Adp31019	Human	sec	198	17	60.7	1170	8	ADP30922	Adp30922	Human	sec
126	17	60.7	555	8	ADP31168	Adp31168	Human	sec	199	17	60.7	1173	8	ADP31155	Adp31155	Human	sec
127	17	60.7	559	8	ADP30716	Adp30716	Human	sec	200	17	60.7	1179	8	ADP30486	Adp30486	Human	sec
128	17	60.7	564	8	ADP30743	Adp30743	Human	sec	201	17	60.7	1189	8	ADP31043	Adp31043	Human	sec
129	17	60.7	567	8	ADP31565	Adp31565	Human	sec	202	17	60.7	1197	8	ADP31342	Adp31342	Human	sec
130	17	60.7	574	8	ADP31133	Adp31133	Human	sec	203	17	60.7	1205	8	ADP30935	Adp30935	Human	sec
131	17	60.7	592	8	ADP30917	Adp30917	Human	sec	204	17	60.7	1252	8	ADP30678	Adp30678	Human	sec
132	17	60.7	624	8	ADP31324	Adp31324	Human	sec	205	17	60.7	1358	8	ADP30995	Adp30995	Human	sec
133	17	60.7	624	8	ADP31325	Adp31325	Human	sec	206	17	60.7	1362	8	ADP31181	Adp31181	Human	sec
134	17	60.7	629	8	ADP31295	Adp31295	Human	sec	207	17	60.7	1383	8	ADP31091	Adp31091	Human	sec
135	17	60.7	645	8	ADP30858	Adp30858	Human	sec	208	17	60.7	1417	8	ADP31160	Adp31160	Human	sec
136	17	60.7	666	8	ADP31256	Adp31256	Human	sec	209	17	60.7	1420	8	ADP30944	Adp30944	Human	sec
137	17	60.7	681	8	ADP30598	Adp30598	Human	sec	210	17	60.7	1435	6	ABU09754	Abu09754	Novel	hum
138	17	60.7	702	8	ADP30868	Adp30868	Human	sec	211	17	60.7	1435	6	ABU90136	Abu90136	Novel	hum
139	17	60.7	702	8	ADP31518	Adp31518	Human	sec	212	17	60.7	1435	6	ABU96438	Abu96438	Novel	hum
140	17	60.7	702	8	ADP31277	Adp31277	Human	sec	213	17	60.7	1435	6	ABU99047	Abu99047	Novel	hum
141	17	60.7	702	8	ADP31278	Adp31278	Human	sec	214	17	60.7	1435	6	ABU98262	Abu98262	Novel	hum
142	17	60.7	708	8	ADP31191	Adp31191	Human	sec	215	17	60.7	1435	6	ABU91868	Abu91868	Novel	hum
143	17	60.7	711	8	ADP31535	Adp31535	Human	sec	216	17	60.7	1435	6	ABU85272	Abu85272	Novel	hum
144	17	60.7	714	8	ADP31440	Adp31440	Human	sec	217	17	60.7	1435	6	ABO00411	AbO00411	Novel	hum
145	17	60.7	714	8	ADP31561	Adp31561	Human	sec	218	17	60.7	1435	6	ABU88962	Abu88962	Novel	hum
146	17	60.7	759	8	ADP31141	Adp31141	Human	sec	219	17	60.7	1435	6	ABO06458	AbO06458	Novel	hum
147	17	60.7	768	8	ADP30912	Adp30912	Human	sec	220	17	60.7	1435	6	ABU95518	Abu95518	Novel	hum
148	17	60.7	771	8	ADP30907	Adp30907	Human	sec	221	17	60.7	1435	6	ABU95208	Abu95208	Novel	hum
149	17	60.7	774	8	ADP30885	Adp30885	Human	sec	222	17	60.7	1435	6	ABU90756	Abu90756	Novel	hum
150	17	60.7	795	8	ADP31332	Adp31332	Human	sec	223	17	60.7	1435	6	ABU93918	Abu93918	Novel	hum
151	17	60.7	799	8	ADP31261	Adp31261	Human	sec	224	17	60.7	1435	6	ABU86192	Abu86192	Novel	hum
152	17	60.7	813	8	ADP30649	Adp30649	Human	sec	225	17	60.7	1435	6	ABU82047	Abu82047	Novel	hum
153	17	60.7	834	8	ADP30738	Adp30738	Human	sec	226	17	60.7	1435	6	ABU07908	Abu07908	Novel	hum
154	17	60.7	843	8	ADP30663	Adp30663	Human	sec	227	17	60.7	1435	6	ABU94228	Abu94228	Novel	hum
155	17	60.7	849	8	ADP31037	Adp31037	Human	sec	228	17	60.7	1435	6	ABEO0101	AbEO0101	Novel	hum
156	17	60.7	849	8	ADP31182	Adp31182	Human	sec	229	17	60.7	1435	6	ABU87112	Abu87112	Novel	hum
157	17	60.7	849	8	ADP31275	Adp31275	Human	sec	230	17	60.7	1435	6	ABU91353	Abu91353	Novel	hum
158	17	60.7	849	8	ADP31276	Adp31276	Human	sec	231	17	60.7	1435	6	ABU90446	Abu90446	Novel	hum
159	17	60.7	861	8	ADP31340	Adp31340	Human	sec	232	17	60.7	1435	6	ABU97037	Abu97037	Novel	hum
160	17	60.7	882	8	ADP30487	Adp30487	Human	sec	233	17	60.7	1435	6	ABO05233	AbO05233	Novel	hum
161	17	60.7	895	8	ADU92573	AdJ92573	Human	PRO	234	17	60.7	1464	6	ADP31437	Adp31437	Human	sec
162	17	60.7	900	6	ADP31337	Adp31337	Human	sec	235	17	60.7	1472	8	ADP31611	Adp31611	Human	sec
163	17	60.7	908	6	ABU08492	Abu08492	Alpha-he1		236	17	60.7	1530	8	ADP31536	Adp31536	Human	sec
164	17	60.7	921	8	ADP31522	Adp31522	Human	sec	237	17	60.7	1550	8	ADP30567	Adp30567	Human	sec
165	17	60.7	930	8	ADP31444	Adp31444	Human	sec	238	17	60.7	1560	8	ADP31631	Adp31631	Human	sec
166	17	60.7	933	8	ADP31140	Adp31140	Human	sec	239	17	60.7	1614	8	ADP31529	Adp31529	Human	sec
167	17	60.7	939	8	ADP31542	Adp31542	Human	sec	240	17	60.7	1647	8	ADP30670	Adp30670	Human	sec
168	17	60.7	939	8	ADP31086	Adp31086	Human	sec	241	17	60.7	1662	8	ADP31419	Adp31419	Human	sec
169	17	60.7	939	8	ADP31541	Adp31541	Human	sec	242	17	60.7	1662	8	ADP31513	Adp31513	Human	sec
170	17	60.7	947	8	ADP30937	Adp30937	Human	sec	243	17	60.7	1700	3	AA18144	AA18144	Plasmodiu	

244	17	60.7	1743	6	ABU88255	Abu88255	Novel	hum	317	17	60.7	3390	8	ADP31148	Adp31148	Human	sec
245	17	60.7	1743	6	ABU90134	Abu90134	Novel	hum	318	17	60.7	3447	8	ADP31112	Adp31112	Human	sec
246	17	60.7	1743	6	ABU96436	Abu96436	Novel	hum	319	17	60.7	4315	5	ADP31308	Adp31308	MUC5B	par
247	17	60.7	1743	6	ABU99045	Abu99045	Novel	hum	320	17	60.7	5304	8	ADP30706	Adp30706	Human	sec
248	17	60.7	1743	6	ABU98260	Abu98260	Novel	hum	321	17	60.7	5514	8	ADP31186	Adp31186	Human	sec
249	17	60.7	1743	6	ABU91966	Abu91966	Novel	hum	322	17	60.7	5514	8	ADP31591	Adp31591	Human	sec
250	17	60.7	1743	6	ABU85270	Abu85270	Novel	hum	323	17	60.7	5703	8	ADL23265	Adl23265	Human	MUC
251	17	60.7	1743	6	ABO00409	AbO00409	Novel	hum	324	17	60.7	5858	9	ABE56507	Aeb56507	Radiochem	
252	17	60.7	1743	6	ABU88960	Abu88960	Novel	hum	325	17	60.7	6465	8	ADP30705	Adp30705	Human	sec
253	17	60.7	1743	6	ABO06456	AbO06456	Novel	hum	326	17	60.7	6729	8	ADP31600	Adp31600	Human	sec
254	17	60.7	1743	6	ABU95516	Abu95516	Novel	hum	327	17	60.7	7285	6	ABJ38280	Abj38280	pAMG21-RA	
255	17	60.7	1743	6	ABU95206	Abu95206	Novel	hum	328	17	60.7	7339	8	AAO16358	Aao16358	Human	tra
256	17	60.7	1743	6	ABU90754	Abu90754	Novel	hum	329	17	60.7	10944	6	ADP31311	Adp31311	Human	sec
257	17	60.7	1743	6	ABU93916	Abu93916	Novel	hum	330	17	60.7	11328	8	ADP31310	Adp31310	Human	sec
258	17	60.7	1743	6	ABU86190	Abu86190	Novel	hum	331	16	57.1	46	9	ADZ47464	Adz47464	Human	HAS
259	17	60.7	1743	6	ABU82045	Abu82045	Novel	hum	332	16	57.1	46	9	AEA34324	Aea34324	Optim pop	
260	17	60.7	1743	6	ABU07906	Abu07906	Novel	hum	333	16	57.1	49	4	AAW24449	Aaw24449	Pig	EST
261	17	60.7	1743	6	ABU94226	Abu94226	Novel	hum	334	16	57.1	51	4	AAU57122	Aau57122	Propionib	
262	17	60.7	1743	6	ABO00099	AbO00099	Novel	hum	335	16	57.1	51	6	ABM53641	Abm53641	Propionib	
263	17	60.7	1743	6	ABU87110	Abu87110	Novel	hum	336	16	57.1	52	4	AAU59153	Aau59153	Propionib	
264	17	60.7	1743	6	ABU91351	Abu91351	Novel	hum	337	16	57.1	52	6	ABM55672	Abm55672	Propionib	
265	17	60.7	1743	6	ABU90444	Abu90444	Novel	hum	338	16	57.1	58	9	ADZ11761	Adz11761	E. coli	D
266	17	60.7	1743	6	ABU97035	Abu97035	Novel	hum	339	16	57.1	85	4	AAU61946	Aau61946	Propionib	
267	17	60.7	1743	6	ABO05231	AbO05231	Novel	hum	340	16	57.1	85	6	ABM58465	Abm58465	Propionib	
268	17	60.7	1776	9	ADY62715	AdY62715	Human	alp	341	16	57.1	94	4	AAAB65963	Aab65963	Human	sec
269	17	60.7	1803	8	ADP31270	Adp31270	Human	sec	342	16	57.1	100	8	ADPX65826	Adpx65826	Plant	ful
270	17	60.7	1803	8	ADP30562	Adp30562	Human	sec	343	16	57.1	109	8	ADP30690	Adp30690	Human	sec
271	17	60.7	1894	6	ABU88253	Abu88253	Novel	hum	344	16	57.1	132	8	ADQ94844	Adq94844	Genetic	b
272	17	60.7	1894	6	ABU90132	Abu90132	Novel	hum	345	16	57.1	133	7	ABO70871	AbO70871	Pseudomon	
273	17	60.7	1894	6	ABU96434	Abu96434	Novel	hum	346	16	57.1	134	8	ADQ94811	Adq94811	Genetic	b
274	17	60.7	1894	6	ABU99043	Abu99043	Novel	hum	347	16	57.1	138	7	ABO67837	AbO67837	Pseudomon	
275	17	60.7	1894	6	ABU98258	Abu98258	Novel	hum	348	16	57.1	142	7	ABO75180	AbO75180	Pseudomon	
276	17	60.7	1894	6	ABU91964	Abu91964	Novel	hum	349	16	57.1	144	7	ABO77101	AbO77101	Pseudomon	
277	17	60.7	1894	6	ABU85258	Abu85258	Novel	hum	350	16	57.1	144	8	ADX72498	Adx72498	Plant	ful
278	17	60.7	1894	6	ABO00407	AbO00407	Novel	hum	351	16	57.1	151	7	ABO81297	AbO81297	Pseudomon	
279	17	60.7	1894	6	ABU88958	Abu88958	Novel	hum	352	16	57.1	154	7	ABO80026	AbO80026	Pseudomon	
280	17	60.7	1894	6	ABO06454	AbO06454	Novel	hum	353	16	57.1	155	7	ABO71527	AbO71527	Pseudomon	
281	17	60.7	1894	6	ABU95514	Abu95514	Novel	hum	354	16	57.1	159	8	ADP30763	Adp30763	Human	sec
282	17	60.7	1894	6	ABU95204	Abu95204	Novel	hum	355	16	57.1	159	8	ADP30762	Adp30762	Human	sec
283	17	60.7	1894	6	ABU90752	Abu90752	Novel	hum	356	16	57.1	160	4	AAW95024	Aaw95024	Human	rep
284	17	60.7	1894	6	ABU93914	Abu93914	Novel	hum	357	16	57.1	160	4	ABB95732	Abb95732	Human	tes
285	17	60.7	1894	6	ABU86188	Abu86188	Novel	hum	358	16	57.1	165	8	ADP31286	Adp31286	Human	sec
286	17	60.7	1894	6	ABU82043	Abu82043	Novel	hum	359	16	57.1	165	8	ADP31174	Adp31174	Human	sec
287	17	60.7	1894	6	ABU07904	Abu07904	Novel	hum	360	16	57.1	171	8	ADP30794	Adp30794	Human	sec
288	17	60.7	1894	6	ABU94224	Abu94224	Novel	hum	361	16	57.1	171	8	ADP31497	Adp31497	Human	sec
289	17	60.7	1894	6	ABO00097	AbO00097	Novel	hum	362	16	57.1	171	8	ADP30793	Adp30793	Human	sec
290	17	60.7	1894	6	ABU87108	Abu87108	Novel	hum	363	16	57.1	180	8	ADP30820	Adp30820	Human	sec
291	17	60.7	1894	6	ABU91349	Abu91349	Novel	hum	364	16	57.1	180	8	ADP30825	Adp30825	Human	sec
292	17	60.7	1894	6	ABU90442	Abu90442	Novel	hum	365	16	57.1	180	8	ADP30828	Adp30828	Human	sec
293	17	60.7	1894	6	ABU97033	Abu97033	Novel	hum	366	16	57.1	180	8	ADP30821	Adp30821	Human	sec
294	17	60.7	1894	6	ABO05229	AbO05229	Novel	hum	367	16	57.1	180	8	ADP30826	Adp30826	Human	sec
295	17	60.7	1968	8	ADP30689	Adp30689	Human	sec	368	16	57.1	180	8	ADP30827	Adp30827	Human	sec
296	17	60.7	2058	8	ADP31630	Adp31630	Human	sec	369	16	57.1	188	8	ADP31660	Adp31660	Human	sec
297	17	60.7	2091	8	ADP31088	Adp31088	Human	sec	370	16	57.1	195	8	ADP30590	Adp30590	Human	sec
298	17	60.7	2109	8	ADT07505	AdT07505	Human	col	371	16	57.1	198	2	AAE59841	Aae59841	AgoR4L	pr
299	17	60.7	2124	8	ADP30560	Adp30560	Human	sec	372	16	57.1	198	8	AAE92113	Aae92113	Human	Apo
300	17	60.7	2182	8	ADP30644	Adp30644	Human	sec	373	16	57.1	198	8	ADP30766	Adp30766	Human	sec
301	17	60.7	2187	8	ADP30882	Adp30882	Human	sec	374	16	57.1	198	8	ADP30767	Adp30767	Human	sec
302	17	60.7	2254	8	ADT07506	AdT07506	Human	col	375	16	57.1	199	7	ABO73562	AbO73562	Pseudomon	
303	17	60.7	2260	8	ADP30687	Adp30687	Human	sec	376	16	57.1	201	8	ADP30711	Adp30711	Human	sec
304	17	60.7	2272	8	ADP31136	Adp31136	Human	sec	377	16	57.1	215	9	ABE803544	Aeb803544	Mycobacte	
305	17	60.7	2272	8	ADP30669	Adp30669	Human	sec	378	16	57.1	215	9	AEA79447	Aea79447	Novel	M.
306	17	60.7	2304	8	ADP31252	Adp31252	Human	sec	379	16	57.1	217	7	ABO72885	AbO72885	Pseudomon	
307	17	60.7	2307	8	ADP31394	Adp31394	Human	sec	380	16	57.1	219	7	ABO76425	AbO76425	Pseudomon	
308	17	60.7	2349	8	ADP30959	Adp30959	Human	sec	381	16	57.1	225	8	ADP30589	Adp30589	Human	sec
309	17	60.7	2382	8	ADP31341	Adp31341	Human	sec	382	16	57.1	228	8	ADP31281	Adp31281	Human	sec
310	17	60.7	2401	8	ADT07507	AdT07507	Human	col	383	16	57.1	229	3	ADCO7930	AdcO7930	Rice	prot
311	17	60.7	2415	8	ADP31023	Adp31023	Human	sec	384	16	57.1	234	3	AAE23898	Aae23898	Arabidops	
312	17	60.7	2418	8	ADP31105	Adp31105	Human	sec	385	16	57.1	234	5	AAE49556	Aae49556	Arabidops	
313	17	60.7	2616	8	ADP31253	Adp31253	Human	sec	386	16	57.1	234	5	AAU93028	Aau93028	Arabidops	
314	17	60.7	2616	9	ABE849675	Aeb849675	N. mening		387	16	57.1	234	7	ADD30781	AdD30781	Plant	yle
315	17	60.7	2976	8	ADP30724	Adp30724	Human	sec	388	16	57.1	234	7	ADL43959	Adl43959	Plant	tra
316	17	60.7	3339	8	ADP31219	Adp31219	Human	sec	389	16	57.1	234	8	ADL61385	Adl61385	A. thalia	

390	16	57.1	234	8	AD002129	Ad002129	Thalactres	463	16	57.1	471	8	ADP30854	ADp30854	Human sec
391	16	57.1	234	8	ADP30484	Adp30484	Human sec	464	16	57.1	472	8	ADP31222	ADp31222	Human sec
392	16	57.1	234	8	ADV09549	Adv09549	A. thalia	465	16	57.1	473	8	ADP31083	Adp31083	Human sec
393	16	57.1	234	8	ADV09550	Adv09550	A. thalia	466	16	57.1	474	8	ADP31235	Adp31235	Human sec
394	16	57.1	243	8	ADP31343	Adp31343	Human sec	467	16	57.1	480	8	ADP31484	Adp31484	Human sec
395	16	57.1	246	8	ADP30619	Adp30619	Human sec	468	16	57.1	489	8	ADP31380	Adp31380	Human sec
396	16	57.1	249	7	AB069987	Ab069987	Pseudomon	469	16	57.1	493	9	ADW69641	Adw69641	HIV-1 gp1
397	16	57.1	254	8	ADP30739	Adp30739	Human sec	470	16	57.1	501	8	ADP31689	Adp31689	Human sec
398	16	57.1	272	8	ADP30693	Adp30693	Human sec	471	16	57.1	510	8	ADP31549	Adp31549	Human sec
399	16	57.1	278	8	ADQ66574	Adq66574	Novel hum	472	16	57.1	519	8	ADP31190	Adp31190	Human sec
400	16	57.1	279	8	ADP31523	Adp31523	Human sec	473	16	57.1	528	8	ADP31525	Adp31525	Human sec
401	16	57.1	279	8	ADP31489	Adp31489	Human sec	474	16	57.1	531	8	ADP31696	Adp31696	Human sec
402	16	57.1	280	7	AB072037	Ab072037	Pseudomon	475	16	57.1	533	8	ADP31279	Adp31279	Human sec
403	16	57.1	285	7	AB072271	Ab072271	Pseudomon	476	16	57.1	543	8	ADP31329	Adp31329	Human sec
404	16	57.1	286	8	ADP30918	Adp30918	Human pro	477	16	57.1	546	8	ADP31449	Adp31449	Human sec
405	16	57.1	291	8	ADP31193	Adp31193	Human sec	478	16	57.1	549	8	ADP30855	Adp30855	Human sec
406	16	57.1	294	8	ADP31045	Adp31045	Human sec	479	16	57.1	549	8	ADP31639	Adp31639	Human sec
407	16	57.1	306	7	AB074423	Ab074423	Pseudomon	480	16	57.1	549	8	ADP31009	Adp31009	Human sec
408	16	57.1	311	2	AAR59843	Aar59843	AporeLx2	481	16	57.1	552	8	ADP31524	Adp31524	Human sec
409	16	57.1	311	2	AAR92115	Aar92115	Human Apo	482	16	57.1	557	8	ADP31103	Adp31103	Human sec
410	16	57.1	315	8	ADP31685	Adp31685	Human sec	483	16	57.1	558	8	ADP31254	Adp31254	Human sec
411	16	57.1	317	8	ADP31615	Adp31615	Human sec	484	16	57.1	558	8	ADP31255	Adp31255	Human sec
412	16	57.1	318	8	ADP30824	Adp30824	Human sec	485	16	57.1	563	8	ADP30881	Adp30881	Human sec
413	16	57.1	318	8	ADP31135	Adp31135	Human sec	486	16	57.1	585	8	ADP31458	Adp31458	Human sec
414	16	57.1	320	7	AB075888	Ab075888	Pseudomon	487	16	57.1	600	8	ADP30865	Adp30865	Human sec
415	16	57.1	331	7	AB070026	Ab070026	Pseudomon	488	16	57.1	604	8	ADP30940	Adp30940	Human sec
416	16	57.1	333	8	ADP30848	Adp30848	Human sec	489	16	57.1	604	8	ADP30941	Adp30941	Human sec
417	16	57.1	336	8	ADP31320	Adp31320	Human sec	490	16	57.1	605	8	ADL12887	Adl12887	Human bre
418	16	57.1	345	8	ADP31016	Adp31016	Human sec	491	16	57.1	610	8	ADP31264	Adp31264	Human sec
419	16	57.1	345	8	ADP31355	Adp31355	Human sec	492	16	57.1	612	8	ADP31169	Adp31169	Human sec
420	16	57.1	346	8	ADK68092	Adk68092	Plant ful	493	16	57.1	618	8	ADP31467	Adp31467	Human sec
421	16	57.1	348	8	ADP30985	Adp30985	Human sec	494	16	57.1	618	8	ADP31466	Adp31466	Human sec
422	16	57.1	348	8	ADP31676	Adp31676	Human sec	495	16	57.1	626	8	ADP30528	Adp30528	Human sec
423	16	57.1	354	8	ADP30783	Adp30783	Human sec	496	16	57.1	627	7	AB068031	Ab068031	Pseudomon
424	16	57.1	354	8	ADP30707	Adp30707	Human sec	497	16	57.1	629	8	ADP30539	Adp30539	Human sec
425	16	57.1	354	8	ADP30779	Adp30779	Human sec	498	16	57.1	633	8	ADP30861	Adp30861	Human sec
426	16	57.1	354	8	ADP30784	Adp30784	Human sec	499	16	57.1	639	8	ADP31265	Adp31265	Human sec
427	16	57.1	354	8	ADP30778	Adp30778	Human sec	500	16	57.1	642	8	ADP30748	Adp30748	Human sec
428	16	57.1	357	8	ADP31560	Adp31560	Human sec	501	16	57.1	642	8	ADP31161	Adp31161	Human sec
429	16	57.1	357	8	ADP30505	Adp30505	Human sec	502	16	57.1	642	8	ADP31124	Adp31124	Human sec
430	16	57.1	360	8	ADP31439	Adp31439	Human sec	503	16	57.1	645	8	ADP31125	Adp31125	Human sec
431	16	57.1	363	8	ADP30659	Adp30659	Human sec	504	16	57.1	645	8	ADP31183	Adp31183	Human sec
432	16	57.1	369	8	ADP30807	Adp30807	Human sec	505	16	57.1	645	8	ADP31022	Adp31022	Human sec
433	16	57.1	372	8	ADP31492	Adp31492	Human sec	506	16	57.1	651	8	ADP31083	Adp31083	Human sec
434	16	57.1	372	8	ADP31610	Adp31610	Human sec	507	16	57.1	658	8	ADP31226	Adp31226	Human sec
435	16	57.1	378	8	ADP31375	Adp31375	Human sec	508	16	57.1	660	8	ADP31606	Adp31606	Human sec
436	16	57.1	381	8	ADP30580	Adp30580	Human sec	509	16	57.1	665	8	ADP30571	Adp30571	Human sec
437	16	57.1	382	8	ADP31221	Adp31221	Human sec	510	16	57.1	666	8	ADP30867	Adp30867	Human sec
438	16	57.1	384	8	ADP31338	Adp31338	Human sec	511	16	57.1	667	8	ADP04463	Adp04463	Sea squit
439	16	57.1	387	8	ADP31368	Adp31368	Human sec	512	16	57.1	669	8	ADP31598	Adp31598	Human sec
440	16	57.1	387	8	ADP31377	Adp31377	Human sec	513	16	57.1	672	8	ADP31667	Adp31667	Human sec
441	16	57.1	392	8	ADT59578	Adt59578	Plant pol	514	16	57.1	672	8	ADP31666	Adp31666	Human sec
442	16	57.1	396	8	ADP31378	Adp31378	Human sec	515	16	57.1	678	8	ADP31585	Adp31585	Human sec
443	16	57.1	399	8	ADP31464	Adp31464	Human sec	516	16	57.1	678	8	ADP31283	Adp31283	Human sec
444	16	57.1	411	7	ABM87022	Abm87022	Rice abio	517	16	57.1	681	8	ADP31053	Adp31053	Human sec
445	16	57.1	411	8	ADP31104	Adp31104	Human sec	518	16	57.1	690	8	ADP30750	Adp30750	Human sec
446	16	57.1	414	8	ADP31477	Adp31477	Human sec	519	16	57.1	699	8	ADP30747	Adp30747	Human sec
447	16	57.1	420	8	ADP31349	Adp31349	Human sec	520	16	57.1	705	8	ADP31330	Adp31330	Human sec
448	16	57.1	421	8	ADP31159	Adp31159	Human sec	521	16	57.1	705	8	ADP31623	Adp31623	Human sec
449	16	57.1	426	8	ADP31331	Adp31331	Human sec	522	16	57.1	710	8	ADP30746	Adp30746	Human sec
450	16	57.1	426	8	ADP31495	Adp31495	Human sec	523	16	57.1	711	8	ADP31652	Adp31652	Human sec
451	16	57.1	429	8	ADP30570	Adp30570	Human sec	524	16	57.1	728	8	ADP30508	Adp30508	Human sec
452	16	57.1	429	8	ADP31207	Adp31207	Human sec	525	16	57.1	729	8	ADP30541	Adp30541	Human sec
453	16	57.1	435	8	ADP31102	Adp31102	Human sec	526	16	57.1	746	8	ADP30992	Adp30992	Human sec
454	16	57.1	438	8	ADP30818	Adp30818	Human sec	527	16	57.1	750	8	ADP30576	Adp30576	Human sec
455	16	57.1	441	8	ADP31206	Adp31206	Human sec	528	16	57.1	750	8	ADP30857	Adp30857	Human sec
456	16	57.1	444	8	ADP31640	Adp31640	Human sec	529	16	57.1	750	8	ADP31131	Adp31131	Human sec
457	16	57.1	449	8	ADQ21283	Adq21283	Human sof	530	16	57.1	750	8	ADP30856	Adp30856	Human sec
458	16	57.1	450	8	ADP31085	Adp31085	Human sec	531	16	57.1	764	8	ADP31614	Adp31614	Human sec
459	16	57.1	453	8	ADP31465	Adp31465	Human sec	532	16	57.1	768	8	ADP31126	Adp31126	Human sec
460	16	57.1	459	8	ADP30932	Adp30932	Human sec	533	16	57.1	772	8	ADP30936	Adp30936	Human sec
461	16	57.1	459	8	ADP30930	Adp30930	Human sec	534	16	57.1	774	8	ADP30506	Adp30506	Human sec
462	16	57.1	467	8	ADX91801	Adx91801	Plant ful	535	16	57.1	774	8	ADP31373	Adp31373	Human sec

536	16	57.1	774	8	ADP30780	Human sec	609	16	57.1	1221	8	ADP30884	Adp30884	Human sec
537	16	57.1	774	8	ADP30781	Human sec	610	16	57.1	1224	8	ADP31426	Adp31426	Human sec
538	16	57.1	774	8	ADP31235	Human sec	611	16	57.1	1227	8	ADP31602	Adp31602	Human sec
539	16	57.1	776	7	ADP71157	Human int	612	16	57.1	1227	8	ADP31210	Adp31210	Human sec
540	16	57.1	776	7	ABM62451	Tumour-as	613	16	57.1	1233	8	ADP30523	Adp30523	Human sec
541	16	57.1	776	9	ADY16233	PRO polyP	614	16	57.1	1239	8	ADP31297	Adp31297	Human sec
542	16	57.1	783	8	ADP31284	Human sec	615	16	57.1	1260	8	ADP30883	Adp30883	Human sec
543	16	57.1	783	8	ADP31436	Human sec	616	16	57.1	1269	8	ADP31189	Adp31189	Human sec
544	16	57.1	792	9	AEA20901	Novel hum	617	16	57.1	1269	8	ADJ23456	Adp23456	Sulfolobu
545	16	57.1	804	8	ADP31635	Human sec	618	16	57.1	1282	8	ADP31328	Adp31328	Human sec
546	16	57.1	807	8	ADP31036	Human sec	619	16	57.1	1302	8	ADP31695	Adp31695	Human sec
547	16	57.1	813	8	ADP31282	Human sec	620	16	57.1	1305	8	ADP31389	Adp31389	Human sec
548	16	57.1	813	8	ADP30561	Human sec	621	16	57.1	1314	8	ADP31197	Adp31197	Human sec
549	16	57.1	816	8	ADP31245	Human sec	622	16	57.1	1344	8	ADP31211	Adp31211	Human sec
550	16	57.1	816	8	ADP31581	Human sec	623	16	57.1	1350	8	ADP31138	Adp31138	Human sec
551	16	57.1	828	8	ADP30878	Human sec	624	16	57.1	1359	8	ADP31129	Adp31129	Human sec
552	16	57.1	831	8	ADP31179	Human sec	625	16	57.1	1365	8	ADP31661	Adp31661	Human sec
553	16	57.1	852	8	ADP30664	Human sec	626	16	57.1	1371	8	ADP31646	Adp31646	Human sec
554	16	57.1	861	8	ADP31020	Human sec	627	16	57.1	1371	8	ADP30876	Adp30876	Human sec
555	16	57.1	861	8	ADP31021	Human sec	628	16	57.1	1380	8	ADP31566	Adp31566	Human sec
556	16	57.1	885	8	ADP31198	Human sec	629	16	57.1	1386	8	ADP31371	Adp31371	Human sec
557	16	57.1	887	8	ADP30554	Human sec	630	16	57.1	1387	8	ADP30946	Adp30946	Human sec
558	16	57.1	887	8	ADP30548	Human sec	631	16	57.1	1431	8	ADP31609	Adp31609	Human sec
559	16	57.1	888	8	ADP30571	Human sec	632	16	57.1	1440	8	ADG34533	Adg34533	Glucocort
560	16	57.1	890	8	ADP31059	Human sec	633	16	57.1	1456	8	ADP30923	Adp30923	Human sec
561	16	57.1	891	8	ADP31668	Human sec	634	16	57.1	1464	8	ADP31040	Adp31040	Human sec
562	16	57.1	897	8	ADP30914	Human sec	635	16	57.1	1476	7	ABO80681	Adp80681	Pseudomon
563	16	57.1	912	8	ADP31507	Human sec	636	16	57.1	1494	8	ADP31650	Adp31650	Human sec
564	16	57.1	918	8	ADP31459	Human sec	637	16	57.1	1506	8	ADP30586	Adp30586	Human sec
565	16	57.1	933	8	ADP31510	Human sec	638	16	57.1	1518	8	ADP31532	Adp31532	Human sec
566	16	57.1	936	8	ADP31486	Human sec	639	16	57.1	1566	8	ADP31407	Adp31407	Human sec
567	16	57.1	936	8	ADP31597	Human sec	640	16	57.1	1578	8	ADP30499	Adp30499	Human sec
568	16	57.1	939	8	ADP30726	Human sec	641	16	57.1	1584	8	ADP31405	Adp31405	Human sec
569	16	57.1	945	8	ADP31238	Human sec	642	16	57.1	1596	8	ADP31491	Adp31491	Human sec
570	16	57.1	945	8	ADP31237	Human sec	643	16	57.1	1602	9	ADZ66459	Adz66459	Human ARN
571	16	57.1	948	8	ADP30586	Human sec	644	16	57.1	1623	8	ADP30552	Adp30552	Human sec
572	16	57.1	966	8	ADP30745	Human sec	645	16	57.1	1623	8	ADP31653	Adp31653	Human sec
573	16	57.1	966	8	ADP31356	Human sec	646	16	57.1	1629	8	ADP30945	Adp30945	Human sec
574	16	57.1	1002	8	ADG39639	Human pan	647	16	57.1	1647	8	ADP31052	Adp31052	Human sec
575	16	57.1	1002	8	ADP30866	Human sec	648	16	57.1	1670	8	ADP30927	Adp30927	Human sec
576	16	57.1	1008	8	ADP30721	Human sec	649	16	57.1	1719	8	ADP31137	Adp31137	Human sec
577	16	57.1	1010	8	ADP31296	Human sec	650	16	57.1	1746	8	ADP30992	Adp30992	Human sec
578	16	57.1	1017	6	AAO26719	SR proteol	651	16	57.1	1749	8	ADP31408	Adp31408	Human sec
579	16	57.1	1017	8	ADP31268	Human sec	652	16	57.1	1771	8	ADP30674	Adp30674	Human sec
580	16	57.1	1023	8	ADP30994	Human sec	653	16	57.1	1782	8	ADP31391	Adp31391	Human sec
581	16	57.1	1030	8	ADP30913	Human sec	654	16	57.1	1789	8	ADP30962	Adp30962	Human sec
582	16	57.1	1050	8	ADP31370	Human sec	655	16	57.1	1794	8	ADP31176	Adp31176	Human sec
583	16	57.1	1056	8	ADP31082	Human sec	656	16	57.1	1815	8	ADP31601	Adp31601	Human sec
584	16	57.1	1057	8	ADP31592	Human sec	657	16	57.1	1827	8	ADP31170	Adp31170	Human sec
585	16	57.1	1062	8	ADP31369	Human sec	658	16	57.1	1848	8	ADP31372	Adp31372	Human sec
586	16	57.1	1065	8	ADP30661	Human sec	659	16	57.1	1875	8	ADP31664	Adp31664	Human sec
587	16	57.1	1065	8	ADP30973	Human sec	660	16	57.1	1878	8	ADP30740	Adp30740	Human sec
588	16	57.1	1065	8	ADP31482	Human sec	661	16	57.1	1917	6	ADA15719	Ada15719	C. elegan
589	16	57.1	1065	8	ADP31287	Human sec	662	16	57.1	1933	8	ADP30889	Adp30889	Human sec
590	16	57.1	1066	8	ADP30563	Human sec	663	16	57.1	1933	8	ADP30902	Adp30902	Human sec
591	16	57.1	1086	8	ADP31447	Human sec	664	16	57.1	1956	8	ADP31662	Adp31662	Human sec
592	16	57.1	1090	8	ADP30187	Human pol	665	16	57.1	1989	8	ADP31185	Adp31185	Human sec
593	16	57.1	1104	8	ADP31156	Human sec	666	16	57.1	1996	8	ADP31599	Adp31599	Human sec
594	16	57.1	1116	8	ADP31692	Human sec	667	16	57.1	2001	8	ADP31644	Adp31644	Human sec
595	16	57.1	1132	8	ADP30952	Human sec	668	16	57.1	2020	8	ADP31056	Adp31056	Human sec
596	16	57.1	1140	8	ADP31128	Human sec	669	16	57.1	2148	8	ADP30974	Adp30974	Human sec
597	16	57.1	1140	8	ADP31130	Human sec	670	16	57.1	2358	8	ADP31690	Adp31690	Human sec
598	16	57.1	1147	8	ADP30965	Human sec	671	16	57.1	2391	8	ADP31366	Adp31366	Human sec
599	16	57.1	1183	8	ADP30536	Human sec	672	16	57.1	2454	8	ADP30469	Adp30469	Human sec
600	16	57.1	1184	7	ADP30653	Human sec	673	16	57.1	2469	8	ADP31031	Adp31031	Human sec
601	16	57.1	1186	2	AAV05840	Banana x1	674	16	57.1	2514	8	ADP30727	Adp30727	Human sec
602	16	57.1	1194	8	ADP31146	Human sec	675	16	57.1	2535	8	ADP31146	Adp31146	Human sec
603	16	57.1	1199	8	ADP31044	Human sec	676	16	57.1	2547	8	ADP31665	Adp31665	Human sec
604	16	57.1	1200	9	ADP31603	Human sec	677	16	57.1	2611	8	ADP31027	Adp31027	Human sec
605	16	57.1	1200	9	ADZ76047	Mouse pre	678	16	57.1	2622	8	ADP31663	Adp31663	Human sec
606	16	57.1	1209	8	ADP30782	Human sec	679	16	57.1	2664	8	ADP31546	Adp31546	Human sec
607	16	57.1	1215	8	ADP30898	Human sec	680	16	57.1	2828	8	ADP30938	Adp30938	Human sec
608	16	57.1	1215	8	ADP30900	Human sec	681	16	57.1	2980	9	AEA36049	Aea36049	Maize Sta

682	16	57.1	3036	8	ADP31595	Adp31595	Human sec	755	15	53.6	111	7	ABO83484	AbO83484	Pseudomon
683	16	57.1	3046	8	ADQ10419	Adq10419	Human pol	756	15	53.6	112	5	ABE89458	ABe89458	Human pol
684	16	57.1	3144	8	ADP31544	Adp31544	Human sec	757	15	53.6	113	8	ADP31617	ADp31617	Human sec
685	16	57.1	3201	8	ADP31545	Adp31545	Human sec	758	15	53.6	118	4	AAU53541	AAu53541	Pseudomon
686	16	57.1	3465	8	ADP31234	Adp31234	Human sec	759	15	53.6	119	6	ABM50060	ABm50060	Pseudomon
687	16	57.1	3477	8	ADP30704	Adp30704	Human sec	760	15	53.6	123	7	ABO70463	ABo70463	Pseudomon
688	16	57.1	3579	8	ADP31098	Adp31098	Human sec	761	15	53.6	125	4	AAU47488	AAu47488	Pseudomon
689	16	57.1	3638	8	ADP30981	Adp30981	Human sec	762	15	53.6	126	6	ABM44007	ABm44007	Pseudomon
690	16	57.1	3907	5	ABG70822	Abg70822	Mouse myo	763	15	53.6	126	6	ABR57124	ABr57124	Camel ant
691	16	57.1	3907	6	ABG74190	Abg74190	Mouse myo	764	15	53.6	127	8	ADP31622	ADp31622	Human sec
692	16	57.1	4752	8	ADP30585	Adp30585	Human sec	765	15	53.6	130	7	ADP27051	ADp27051	Human adi
693	16	57.1	4752	8	ADP30651	Adp30651	Human sec	766	15	53.6	131	7	ABO66182	ABo66182	Klebsiell
694	16	57.1	5397	8	ADP31068	Adp31068	Human sec	767	15	53.6	134	7	ABO76972	ABo76972	Human sec
695	16	57.1	8973	8	ADP31119	Adp31119	Human sec	768	15	53.6	135	7	ABO83194	ABo83194	Pseudomon
696	15	53.6	12	6	ABG71386	Abg71386	Staphyloc	769	15	53.6	135	7	ABO79475	ABo79475	Pseudomon
697	15	53.6	24	8	ADD95109	Add95109	PCR prime	770	15	53.6	135	8	ADP30703	ADp30703	Human sec
698	15	53.6	24	8	ADU56677	Adu56677	GSP-4 c10	771	15	53.6	136	7	ABO74222	ABo74222	Pseudomon
699	15	53.6	27	4	AAU13578	AAu13578	Peptide #	772	15	53.6	138	3	AAAB42027	AAa42027	Human ORF
700	15	53.6	27	4	ABR32506	Abb32506	Peptide #	773	15	53.6	139	7	ABO68581	ABo68581	Pseudomon
701	15	53.6	27	4	AAU25975	AAu25975	Peptide #	774	15	53.6	142	8	ADP31632	ADp31632	Human sec
702	15	53.6	27	4	ABR27361	Abb27361	Human pep	775	15	53.6	143	7	ABO77067	ABo77067	Pseudomon
703	15	53.6	27	4	ABR18013	Abb18013	Protein #	776	15	53.6	144	7	ABO75730	ABo75730	Pseudomon
704	15	53.6	27	4	AAU65718	AAu65718	Human bon	777	15	53.6	146	4	ABG02814	ABg02814	Novel hum
705	15	53.6	27	4	AAU53341	AAu53341	Human bra	778	15	53.6	150	8	AD145261	AD145261	Rice isop
706	15	53.6	27	4	ABG47359	Abg47359	Human liv	779	15	53.6	152	8	ADP31605	ADp31605	Human sec
707	15	53.6	27	4	AAU01330	AAu01330	Peptide #	780	15	53.6	153	7	ABO83454	ABo83454	Pseudomon
708	15	53.6	27	5	ABG35348	Abg35348	Human pep	781	15	53.6	154	7	ABO75839	ABo75839	Pseudomon
709	15	53.6	34	4	AAU17951	AAu17951	Peptide #	782	15	53.6	156	7	ABO75667	ABo75667	Pseudomon
710	15	53.6	34	4	ABR36982	Abb36982	Peptide #	783	15	53.6	158	3	AAU29118	AAg29118	Arabidops
711	15	53.6	41	8	ADP21826	Adp21826	Low densi	784	15	53.6	163	7	ADU78226	ADU78226	Human CGD
712	15	53.6	61	4	AAU51383	AAu51383	Pseudomon	785	15	53.6	163	7	ABO80383	ABo80383	Pseudomon
713	15	53.6	61	6	ABU47992	ABu47992	Pseudomon	786	15	53.6	165	8	ADP30587	ADp30587	Human sec
714	15	53.6	69	6	ABU62232	ABu62232	Pseudomon	787	15	53.6	165	8	ADP31041	ADp31041	Human sec
715	15	53.6	69	6	ABM58751	ABm58751	Pseudomon	788	15	53.6	168	8	ADP30841	ADp30841	Human sec
716	15	53.6	70	6	ABU04032	ABu04032	Human mus	789	15	53.6	169	7	ABO74424	ABo74424	Pseudomon
717	15	53.6	70	6	ABU13336	ABu13336	Novel hum	790	15	53.6	169	7	ABO70973	ABo70973	Pseudomon
718	15	53.6	72	7	ABO66669	ABo6669	Klebsiell	791	15	53.6	171	8	ADP30599	ADp30599	Human sec
719	15	53.6	72	7	ABO66669	ABo6669	Klebsiell	792	15	53.6	173	8	ADY10989	ADy10989	Plant ful
720	15	53.6	76	4	ABR37637	Abb37637	Peptide #	793	15	53.6	174	8	ADP31530	ADp31530	Human sec
721	15	53.6	76	4	AAU70732	AAu70732	Human bon	794	15	53.6	177	7	ABO70036	ABo70036	Pseudomon
722	15	53.6	76	4	ABG52439	ABg52439	Human liv	795	15	53.6	177	8	ADP30755	ADp30755	Human sec
723	15	53.6	76	4	AAU06152	AAu06152	Peptide #	796	15	53.6	179	7	ABO78051	ABo78051	Pseudomon
724	15	53.6	77	4	AAU91132	AAu91132	Human imm	797	15	53.6	183	4	AAU54675	AAU54675	Pseudomon
725	15	53.6	80	4	AAU59349	AAu59349	Pseudomon	798	15	53.6	183	6	ABM51194	ABM51194	Pseudomon
726	15	53.6	80	6	ABM55868	ABm55868	Pseudomon	799	15	53.6	184	7	ABO77341	ABo77341	Pseudomon
727	15	53.6	80	7	ABO71261	ABo71261	Pseudomon	800	15	53.6	187	9	ABE03464	ABe03464	Mycobacte
728	15	53.6	83	4	AAU53165	AAu53165	Pseudomon	801	15	53.6	187	9	AEA79367	AEa79367	Novel M.
729	15	53.6	83	6	ABM49684	ABm49684	Pseudomon	802	15	53.6	188	6	ABM69571	ABM69571	Phototrab
730	15	53.6	85	4	AAU46688	AAu46688	Pseudomon	803	15	53.6	188	7	ABO81563	ABo81563	Pseudomon
731	15	53.6	85	6	ABM43387	ABm43387	Pseudomon	804	15	53.6	189	8	ADP31641	ADp31641	Human sec
732	15	53.6	86	7	ABO78259	ABo78259	Pseudomon	805	15	53.6	189	8	ADP30641	ADp30641	Human sec
733	15	53.6	87	4	AAU41562	AAu41562	Pseudomon	806	15	53.6	192	7	AD163118	AD163118	Human epo
734	15	53.6	87	4	ABG07613	ABg07613	Novel hum	807	15	53.6	192	8	ADP30573	ADp30573	Human sec
735	15	53.6	87	6	ABM38081	ABm38081	Pseudomon	808	15	53.6	192	8	ADP30575	ADp30575	Human sec
736	15	53.6	88	7	ABO76533	ABo76533	Pseudomon	809	15	53.6	192	8	ADU73861	ADU73861	Plant ful
737	15	53.6	89	4	ABG27446	ABg27446	Novel hum	810	15	53.6	194	7	ABO78166	ABo78166	Pseudomon
738	15	53.6	90	6	ADK34238	ADk34238	Actinecto	811	15	53.6	194	8	ADP31626	ADp31626	Human sec
739	15	53.6	95	4	AAU50672	AAu50672	Pseudomon	812	15	53.6	195	8	ADP30696	ADp30696	Human sec
740	15	53.6	95	4	AAU40381	AAu40381	Pseudomon	813	15	53.6	198	8	ADP30493	ADp30493	Human sec
741	15	53.6	95	6	ABM47191	ABm47191	Pseudomon	814	15	53.6	198	8	ADP30492	ADp30492	Human sec
742	15	53.6	95	6	ABM36900	ABm36900	Pseudomon	815	15	53.6	198	8	ADP30477	ADp30477	Human sec
743	15	53.6	101	4	AAU83737	AAu83737	Human imm	816	15	53.6	198	8	ADP30481	ADp30481	Human sec
744	15	53.6	102	4	AAU61573	AAu61573	Pseudomon	817	15	53.6	200	7	ABO73557	ABo73557	Pseudomon
745	15	53.6	102	6	ABM58092	ABm58092	Pseudomon	818	15	53.6	200	7	ABO73557	ABo73557	Pseudomon
746	15	53.6	103	6	AAU48984	AAu48984	Pseudomon	819	15	53.6	201	3	ADP30686	ADp30686	Arabidops
747	15	53.6	103	6	ABM45503	ABm45503	Pseudomon	820	15	53.6	204	8	ADP30488	ADp30488	Human sec
748	15	53.6	103	7	ABO84304	ABo84304	Pseudomon	821	15	53.6	204	8	ADP31421	ADp31421	Human sec
749	15	53.6	106	2	AAU29187	AAu29187	Amino aci	822	15	53.6	204	8	ADP31422	ADp31422	Human sec
750	15	53.6	107	4	AAU43314	AAu43314	Pseudomon	823	15	53.6	204	8	ADP30489	ADp30489	Human sec
751	15	53.6	107	5	ABP64480	ABp64480	Human ORF	824	15	53.6	204	8	ADP31424	ADp31424	Human sec
752	15	53.6	107	5	ABM39833	ABm39833	Pseudomon	825	15	53.6	204	8	ADP31420	ADp31420	Human sec
753	15	53.6	108	3	AAU21190	AAu21190	Exo14 par	826	15	53.6	204	8	ADP30490	ADp30490	Human sec
754	15	53.6	110	9	ABM95190	ABm95190	M. xanthu	827	15	53.6	204	8	ADP30490	ADp30490	Human sec


```
974 15 53.6 587 4 ABB67608 Abb67608 Drosophil
975 15 53.6 588 8 ADP30877 Adp30877 Human sec
976 15 53.6 588 8 ADP31680 Adp31680 Human sec
977 15 53.6 588 8 ADP31699 Adp31699 Human sec
978 15 53.6 591 2 AAY15228 Aay15228 Human rec
979 15 53.6 594 8 ADP30895 Adp30895 Human sec
980 15 53.6 603 8 ADP31150 Adp31150 Human sec
981 15 53.6 603 8 ADP31645 Adp31645 Human sec
982 15 53.6 605 8 ADP30507 Adp30507 Human sec
983 15 53.6 608 8 ADP04735 Adp04735 Sea squit
984 15 53.6 609 8 ADP31263 Adp31263 Human sec
985 15 53.6 612 7 AB069512 Ab069512 Pseudomon
986 15 53.6 612 8 ADP31064 Adp31064 Human sec
987 15 53.6 615 8 ADP31361 Adp31361 Human sec
988 15 53.6 615 8 ADP30803 Adp30803 Human sec
989 15 53.6 615 8 ADP31360 Adp31360 Human sec
990 15 53.6 617 8 ADP31657 Adp31657 Human sec
991 15 53.6 621 8 ADP30896 Adp30896 Human sec
992 15 53.6 621 8 ADP31147 Adp31147 Human sec
993 15 53.6 632 7 AB076798 Ab076798 Pseudomon
994 15 53.6 637 8 ADP31397 Adp31397 Human sec
995 15 53.6 638 8 ADP30513 Adp30513 Human sec
996 15 53.6 640 8 ADP30977 Adp30977 Human sec
997 15 53.6 642 8 ADP30977 Adp30977 Human sec
998 15 53.6 645 8 ADP31067 Adp31067 Human sec
999 15 53.6 648 8 ADP31306 Adp31306 Human sec
1000 15 53.6 651 8 ADP31543 Adp31543 Human sec
```

ALIGNMENTS

```
RESULT 1
AB082562
ID AB082562 standard; protein; 104 AA.
XX
AC AB082562;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #14737.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, DeLoughery C, Bush D;
XX
DR WPI: 2003-615309/58.
XX
DR N-PSDB: ABD16133.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 31308; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
XX
CC polynucleotides encoding them. The sequences are useful in diagnosis and
XX
CC therapy of pathological conditions, as molecular targets for diagnostics,
XX
CC prophylaxis and treatment of pathological conditions resulting from a
XX
CC bacterial infection, for evaluating a compound, such as a polypeptide,
```

```
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences AB067826-
CC AB084336 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC segdata.uspto.gov/sequence.html
XX
SQ Sequence 104 AA;
XX
Query Match 64.3%; Score 18; DB 7; Length 104;
Best Local Similarity 16.7%; Pred. No. 4.9e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXC 12
Db 56 CSATASASTTC 67
XX
RESULT 2
ADP30723
ID ADP30723 standard; protein; 110 AA.
XX
AC ADP30723;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1490.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX
PR 29-AUG-2002; 2002US-0406579P.
XX
PR 29-AUG-2002; 2002US-0406585P.
XX
PR 29-AUG-2002; 2002US-0406588P.
XX
PR 29-AUG-2002; 2002US-0406608P.
XX
PR 29-AUG-2002; 2002US-0406611P.
XX
PR 29-AUG-2002; 2002US-0406612P.
XX
PR 29-AUG-2002; 2002US-0406616P.
XX
PR 29-AUG-2002; 2002US-0406640P.
XX
PR 29-AUG-2002; 2002US-0406642P.
XX
PR 29-AUG-2002; 2002US-0406646P.
XX
PR 29-AUG-2002; 2002US-0406653P.
XX
PR 29-AUG-2002; 2002US-0406655P.
XX
PR 29-AUG-2002; 2002US-0406666P.
XX
PR 17-SEP-2002; 2002US-0410946P.
XX
PR 17-SEP-2002; 2002US-0410947P.
XX
PR 17-SEP-2002; 2002US-0410948P.
XX
PR 17-SEP-2002; 2002US-0410949P.
XX
PR 17-SEP-2002; 2002US-0410953P.
XX
PR 17-SEP-2002; 2002US-0410957P.
XX
PR 17-SEP-2002; 2002US-0410958P.
XX
PR 17-SEP-2002; 2002US-0410959P.
XX
PR 17-SEP-2002; 2002US-0410960P.
XX
PR 17-SEP-2002; 2002US-0410961P.
XX
PR 17-SEP-2002; 2002US-0410962P.
XX
PR 17-SEP-2002; 2002US-0411019P.
XX
PR 17-SEP-2002; 2002US-0411022P.
XX
PR 17-SEP-2002; 2002US-0411023P.
XX
PR 17-SEP-2002; 2002US-0411024P.
```

PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-047136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 09-JUN-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MW, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX MPI; 2004-348438/32.
XX
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 2721; 428bp; English.
XX
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
XX
SQ Sequence 110 AA;

Query Match 64.3%; Score 18; DB 8; Length 110;
Best Local Similarity 16.7%; Pred. No. 5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 93 CATMAATATAAC 104

RESULT 3
ADP31638
ID ADP31638 standard; protein; 135 AA.
XX
XX ADP31638;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2405.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX
XX 28-AUG-2003; 2003MO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411039P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411089P.
XX 17-SEP-2002; 2002US-0411101P.
XX 17-SEP-2002; 2002US-0411111P.
XX 18-APR-2003; 2003US-0463700P.
XX 18-APR-2003; 2003US-0463708P.
XX 18-APR-2003; 2003US-0463716P.
XX 18-APR-2003; 2003US-0463732P.
XX 02-MAY-2003; 2003US-0467199P.
XX 02-MAY-2003; 2003US-0467201P.
XX 02-MAY-2003; 2003US-0467203P.
XX 02-MAY-2003; 2003US-0467230P.
XX 19-MAY-2003; 2003US-0471306P.

PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 22030; 455bp; English.
XX
CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-*P. aeruginosa* drugs, as templates for recombinant
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology. Sequences AB067826-
CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 178 AA;
XX
Query Match 64.3%; Score 18; DB 7; Length 178;
Best Local Similarity 16.7%; Pred. No. 5, 5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXXC 12
Db 92 CATATAASTSC 103
XX
RESULT 6
ADP31527
ID ADP31527 standard; protein; 264 AA.
XX
AC ADP31527;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2294.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX
PR 29-AUG-2002; 2002US-0406579P.
XX
PR 29-AUG-2002; 2002US-0406585P.
XX
PR 29-AUG-2002; 2002US-0406588P.
XX
PR 29-AUG-2002; 2002US-0406608P.
XX
PR 29-AUG-2002; 2002US-0406611P.
XX
PR 29-AUG-2002; 2002US-0406612P.
XX
PR 29-AUG-2002; 2002US-0406616P.
XX
PR 29-AUG-2002; 2002US-0406640P.
XX
PR 29-AUG-2002; 2002US-0406642P.
XX
PR 29-AUG-2002; 2002US-0406646P.
XX
PR 29-AUG-2002; 2002US-0406653P.
XX
PR 29-AUG-2002; 2002US-0406655P.
XX
PR 29-AUG-2002; 2002US-0406666P.
XX
PR 17-SEP-2002; 2002US-0410946P.
XX
PR 17-SEP-2002; 2002US-0410947P.
XX
PR 17-SEP-2002; 2002US-0410948P.
XX
PR 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0467133P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485255P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halembeck RF, Huang MM, Kothakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1, SEQ ID NO 3525; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.

```
XX SQ Sequence 264 AA;
Query Match 64.3%; Score 18; DB 8; Length 264;
Best Local Similarity 16.7%; Pred. No. 6e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 12
Db 109 CAAATATTTTC 120

RESULT 7
ADP31435
ID ADP31435 standard; protein; 270 AA.
XX
XX ADP31435;
AC 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #2202.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX
XX 29-AUG-2002; 2002US-0406579P.
XX
XX 29-AUG-2002; 2002US-0406585P.
XX
XX 29-AUG-2002; 2002US-0406588P.
XX
XX 29-AUG-2002; 2002US-0406608P.
XX
XX 29-AUG-2002; 2002US-0406611P.
XX
XX 29-AUG-2002; 2002US-0406612P.
XX
XX 29-AUG-2002; 2002US-0406616P.
XX
XX 29-AUG-2002; 2002US-0406642P.
XX
XX 29-AUG-2002; 2002US-0406646P.
XX
XX 29-AUG-2002; 2002US-0406653P.
XX
XX 29-AUG-2002; 2002US-0406655P.
XX
XX 29-AUG-2002; 2002US-0406666P.
XX
XX 17-SEP-2002; 2002US-0410946P.
XX
XX 17-SEP-2002; 2002US-0410947P.
XX
XX 17-SEP-2002; 2002US-0410948P.
XX
XX 17-SEP-2002; 2002US-0410949P.
XX
XX 17-SEP-2002; 2002US-0410953P.
XX
XX 17-SEP-2002; 2002US-0410957P.
XX
XX 17-SEP-2002; 2002US-0410958P.
XX
XX 17-SEP-2002; 2002US-0410959P.
XX
XX 17-SEP-2002; 2002US-0410960P.
XX
XX 17-SEP-2002; 2002US-0410961P.
XX
XX 17-SEP-2002; 2002US-0410962P.
XX
XX 17-SEP-2002; 2002US-0411019P.
XX
XX 17-SEP-2002; 2002US-0411022P.
XX
XX 17-SEP-2002; 2002US-0411023P.
XX
XX 17-SEP-2002; 2002US-0411024P.
XX
XX 17-SEP-2002; 2002US-0411032P.
XX
XX 17-SEP-2002; 2002US-0411035P.
XX
XX 17-SEP-2002; 2002US-0411037P.
XX
XX 17-SEP-2002; 2002US-0411041P.
XX
XX 17-SEP-2002; 2002US-0411045P.
XX
XX 17-SEP-2002; 2002US-0411046P.
XX
XX 17-SEP-2002; 2002US-0411048P.
XX
XX 17-SEP-2002; 2002US-0411052P.
XX
XX 17-SEP-2002; 2002US-0411055P.
XX
XX 17-SEP-2002; 2002US-0411073P.
XX
XX 17-SEP-2002; 2002US-0411082P.
```

```
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493572P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong UG, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX PS Claim 1, SEQ ID NO 3433; 428bp; English.
XX
XX CC The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPOMEB and is not in the specification.
XX
XX SQ Sequence 270 AA;
Query Match 64.3%; Score 18; DB 8; Length 270;
Best Local Similarity 16.7%; Pred. No. 6e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 12
Db 179 CTATATTAAAC 190

RESULT 8
ABB65879
ID ABB65879 standard; protein; 307 AA.
XX
XX ABB65879;
AC 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 24429.
XX
```


KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO20017042-A2.
XX
PD 27-SEP-2001.
XX
PP 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL09982.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 24429; 21pp + Sequence listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16175-AB130511), expressed DNA
CC sequences (AB161840-AB16175) and the encoded proteins (AB5737-
CC AB872072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 307 AA;
Query Match 64.3%; Score 18; DB 4; Length 307;
Best Local Similarity 16.7%; Pred. No. 6.2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXC 12
Db 44 CTTTTTTTTTTC 55
RESULT 9
ADP31223
ID ADP31223 standard; protein; 357 AA.
XX
XX ADP31223;
AC
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1990.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411025P.
PR 17-SEP-2002; 2002US-0411033P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486891P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493572P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LR, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Hang MM, Kothakota S, Hsieh L, Linemann T;
PI Pierce K, Wang Y, Wong JGF, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PR such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3221; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
SQ Sequence 357 AA;

Query Match 64.3%; Score 18; DB 8; Length 357;
Best Local Similarity 16.7%; Pred. No. 6.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 175 CTTAATTTC 186

RESULT 10
ADP31132
ID ADP31132 standard; protein; 615 AA.
XX
AC ADP31132;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1899.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411029P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493373P.
PR 08-AUG-2003; 2003US-0493577P.

XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
XX Halenbeck RF, Huang MM, Kothakota S, Haislan L, Linnemann T;
XX Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3130; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPWEB and is not in the specification.
SQ Sequence 615 AA;

Query Match 64.3%; Score 18; DB 8; Length 615;
Best Local Similarity 16.7%; Pred. No. 7.1e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
Db 175 CATTATTTTAAAC 186

RESULT 11

AAU51580
ID AAU51580 standard; protein; 627 AA.

AAU51580;

27-FEB-2002 (first entry)

Propionibacterium acnes immunogenic protein #12476.

SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis; uveitis; endophthalmitis; bone; joint; central nervous system; ELISA, inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay; dermatological; osteopathic; neuroprotectant.

Propionibacterium acnes.

MO200181581-A2.

01-NOV-2001.

20-APR-2001; 2001WO-US012865.

21-APR-2000; 2000US-0199047P.

02-JUN-2000; 2000US-0208841P.

07-JUL-2000; 2000US-0216747P.

(CORI-) CORIXA CORP.

Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhattacharya A;

L'maisonneuve J, Zhang Y, Jen S, Carter D;

WPI; 2001-616774/71.

N-PSDB; AAS59551.

Example 1; SEQ ID NO 12775; 1069pp; English.

Sequences AAU5105-AAU68017 represent Propionibacterium acnes immunogenic

polypeptides. The proteins and their associated DNA sequences are used in

the treatment, prevention and diagnosis of medical conditions caused by

P. acnes. The disorders include SAPHO syndrome (synovitis, acne,

pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.

P. acnes is also involved in infections of bone, joints and the central

nervous system, however it is particularly involved in the inflammatory

lesions associated with acne vulgaris. A method for detecting the

presence or absence of P. acnes in a patient comprises contacting a

sample with a binding agent that binds to the proteins of the invention

and determining the amount of bound protein in the sample. The

polypeptides may be used as antigens in the production of antibodies

specific for P. acnes proteins. These antibodies can be used to

downregulate expression and activity of P. acnes polypeptides and

therefore treat P. acnes infections. The antibodies may also be used as

diagnostic agents for determining P. acnes presence, for example, by

enzyme linked immunosorbent assay (ELISA). Note: The sequence data for

this patent did not form part of the printed specification, but was

obtained in electronic format directly from WIPO at

ftp.wipo.int/pub/published_pct_sequences

Sequence 627 AA;

Query Match 64.3%; Score 18; DB 4; Length 627;

Best Local Similarity 16.7%; Pred. No. 7.2e-05;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
Db 46 CATTTSATNSC 57

RESULT 12

ABM48099
ID ABM48099 standard; protein; 627 AA.

ABM48099;

20-OCT-2003 (first entry)

Propionibacterium acnes predicted ORF-encoded polypeptide #12775.

Acne vulgaris; antisephorhoic; dermatological; antibacterial;
immunostimulant; immune response; vaccine.

Propionibacterium acnes.

MO2003033515-A1.

24-APR-2003.

11-OCT-2002; 2002WO-US032727.

15-OCT-2001; 2001US-00978825.

(CORI-) CORIXA CORP.

Mitcham JL, Skeiky YAW, Persing DH, Bhattacharya A, Maisonneuve JL;

Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

Barth B, Vallieve-Douglas J;

WPI; 2003-381789/36.

N-PSDB; ACF64480.

Example 1; SEQ ID NO 12775; 1481pp; English.

The invention relates to an isolated polynucleotide (ACF64435-ACF64733)

encoding a Propionibacterium acnes protein. The invention also relates to

polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to

immunogenic fragments of P. acnes polypeptides. The invention

additionally encompasses expression vectors and host cells comprising a

polynucleotide of the invention; antibodies against polypeptides of the

invention; fusion proteins comprising a polypeptide of the invention; a

method for stimulating an immune response specific for a P. acnes

polypeptide and an isolated T cell population comprising T cells prepared

via this method; a vaccine composition (comprising P. acnes polypeptides,

polynucleotides, antibodies, fusion proteins, T cell populations, or

antigen-presenting cells that express the polypeptide); a method and kit

for detecting or determining the presence or absence of P. acnes in a

patient; and a method for inhibiting the development of P. acnes in a

patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion

proteins, T cell populations or antigen-presenting cells that express the

polypeptides are useful for diagnosing, preventing or treating acne

vulgaris, or for stimulating an immune response specific for a P. acnes

protein. The polynucleotides can also be used as probes or primers for

nucleic acid hybridization. The vaccine composition is useful for the

stimulation of an immune response against P. acnes, or for treating acne,

and the kit is useful for performing a diagnostic assay. The present

sequence represents a polypeptide predicted to be encoded by an ORF (open

reading frame) contained within the P. acnes polynucleotides of the

invention. Note: The sequence data for this patent did not form part of

the printed specification, but was obtained in electronic format directly

from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 627 AA;

Query Match 64.3%; Score 18; DB 4; Length 627;

Best Local Similarity 16.7%; Pred. No. 7.2e-05;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Query Match          64.3%; Score 18; DB 6; Length 627;
Best Local Similarity 16.7%; Pred. No. 7.2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      1 CXXXXXXXXXXC 12
DB      46 CAATTTTATSC 57

RESULT 13
ADP31142
ID ADP31142 standard; protein; 669 AA.
XX
AC ADP31142;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1909.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN W02004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX
PR 29-AUG-2002; 2002US-0406579P.
XX
PR 29-AUG-2002; 2002US-0406585P.
XX
PR 29-AUG-2002; 2002US-0406588P.
XX
PR 29-AUG-2002; 2002US-0406608P.
XX
PR 29-AUG-2002; 2002US-0406611P.
XX
PR 29-AUG-2002; 2002US-0406612P.
XX
PR 29-AUG-2002; 2002US-0406616P.
XX
PR 29-AUG-2002; 2002US-0406640P.
XX
PR 29-AUG-2002; 2002US-0406642P.
XX
PR 29-AUG-2002; 2002US-0406646P.
XX
PR 29-AUG-2002; 2002US-0406653P.
XX
PR 29-AUG-2002; 2002US-0406655P.
XX
PR 29-AUG-2002; 2002US-0406666P.
XX
PR 17-SEP-2002; 2002US-0410946P.
XX
PR 17-SEP-2002; 2002US-0410947P.
XX
PR 17-SEP-2002; 2002US-0410948P.
XX
PR 17-SEP-2002; 2002US-0410949P.
XX
PR 17-SEP-2002; 2002US-0410953P.
XX
PR 17-SEP-2002; 2002US-0410957P.
XX
PR 17-SEP-2002; 2002US-0410958P.
XX
PR 17-SEP-2002; 2002US-0410959P.
XX
PR 17-SEP-2002; 2002US-0410960P.
XX
PR 17-SEP-2002; 2002US-0410961P.
XX
PR 17-SEP-2002; 2002US-0410962P.
XX
PR 17-SEP-2002; 2002US-0411019P.
XX
PR 17-SEP-2002; 2002US-0411022P.
XX
PR 17-SEP-2002; 2002US-0411023P.
XX
PR 17-SEP-2002; 2002US-0411024P.
XX
PR 17-SEP-2002; 2002US-0411032P.
XX
PR 17-SEP-2002; 2002US-0411035P.
XX
PR 17-SEP-2002; 2002US-0411037P.
XX
PR 17-SEP-2002; 2002US-0411041P.
XX
PR 17-SEP-2002; 2002US-0411045P.
XX
PR 17-SEP-2002; 2002US-0411046P.
XX
PR 17-SEP-2002; 2002US-0411048P.
XX
PR 17-SEP-2002; 2002US-0411052P.
XX
PR 17-SEP-2002; 2002US-0411055P.
XX
PR 17-SEP-2002; 2002US-0411073P.
XX
PR 17-SEP-2002; 2002US-0411082P.
XX
PR 17-SEP-2002; 2002US-0411101P.
XX
PR 17-SEP-2002; 2002US-0411111P.
XX
PR 18-APR-2003; 2003US-0463700P.
```

```
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471366P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486909P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Halahan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3140; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPOWEB and is not in the specification.
XX
XX Sequence 669 AA;
SQ
Query Match          64.3%; Score 18; DB 8; Length 669;
Best Local Similarity 16.7%; Pred. No. 7.3e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      1 CXXXXXXXXXXC 12
DB      652 CAATTTTAAC 663

RESULT 14
ADP31215
ID ADP31215 standard; protein; 711 AA.
XX
AC ADP31215;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1982.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
```

```
OS Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 29-AUG-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411101P.
XX 17-SEP-2002; 2002US-0411111P.
XX 18-APR-2003; 2003US-0463700P.
XX 18-APR-2003; 2003US-0463708P.
XX 18-APR-2003; 2003US-0463716P.
XX 18-APR-2003; 2003US-0463732P.
XX 02-MAY-2003; 2003US-0467199P.
XX 02-MAY-2003; 2003US-0467201P.
XX 02-MAY-2003; 2003US-0467203P.
XX 02-MAY-2003; 2003US-0467230P.
XX 19-MAY-2003; 2003US-0471306P.
XX 19-MAY-2003; 2003US-047136P.
XX 22-MAY-2003; 2003US-0472420P.
XX 22-MAY-2003; 2003US-0472430P.
XX 09-JUN-2003; 2003US-0476609P.
XX 09-JUN-2003; 2003US-0476641P.
XX 08-JUL-2003; 2003US-0485218P.
XX 08-JUL-2003; 2003US-0485233P.
XX 08-JUL-2003; 2003US-048524P.
XX 08-JUL-2003; 2003US-0485325P.
XX 14-JUL-2003; 2003US-0486446P.
XX 14-JUL-2003; 2003US-0486480P.
XX 15-JUL-2003; 2003US-0486891P.
XX 15-JUL-2003; 2003US-0486960P.
XX 08-AUG-2003; 2003US-0493341P.
XX
XX 08-AUG-2003; 2003US-0493370P.
XX 08-AUG-2003; 2003US-0493573P.
XX 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
XX PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
XX PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI, 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX PT genetic, bacterial and viral diseases.
XX
XX Claim 1, SEQ ID NO 3213; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cytostatic,
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPWEB and is not in the specification.
XX
XX SEQ Sequence 711 AA;
XX
XX Query Match 64.3%; Score 18; DB 8; Length 711;
XX Best Local Similarity 16.7%; Pred. No. 7.4e-05;
XX Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
XX 1 CXXXXXXXKX 12
XX Db 684 CAAAAAATTAC 695
XX
XX RESULT 15
XX ADP31196
XX ID ADP31196 standard; protein; 739 AA.
XX
XX ADP31196;
XX AC
XX ADP31196;
XX
XX 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #1963.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
```

PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485244P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486691P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3194; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and

CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 739 AA;
Query Match 64.3%; Score 18; DB 8; Length 739;
Best Local Similarity 16.7%; Pred. No. 7.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXXC 12
DB 226 CAAAAAAAAAAC 237
RESULT 16
ADP31244
ID ADP31244 standard; protein; 771 AA.
XX
XX ADP31244;
AC XX
XX
DT 12-AUG-2004 (first entry)
DT XX
DE Human secreted protein SEQ ID #2011.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
PD 29-APR-2004.
XX
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0410963P.
XX 17-SEP-2002; 2002US-0410964P.
XX 17-SEP-2002; 2002US-0410965P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411033P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.

PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486891P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halembeck RF, Huang MM, Kothakota S, Haislan L, Linemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
XX Claim 1; SEQ ID NO 3218; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
XX Sequence 876 AA;
SQ
Query Match 64.3%; Score 18; DB 8; Length 876;
Best Local Similarity 16.7%; Pred. No. 7.7e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXC 12
Db 28 CATATATAAATC 39
RESULT 18
ADP31688
ID ADP31688 standard; protein; 882 AA.
XX
XX ADP31688;
AC
XX
XX 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #245.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX
XX 29-APR-2004.
XX
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX
XX 29-AUG-2002; 2002US-0406576P.
XX
XX 29-AUG-2002; 2002US-0406579P.
XX
XX 29-AUG-2002; 2002US-0406585P.
XX
XX 29-AUG-2002; 2002US-0406588P.
XX
XX 29-AUG-2002; 2002US-0406608P.
XX
XX 29-AUG-2002; 2002US-0406611P.
XX
XX 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406665P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463715P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476611P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486891P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halembeck RF, Huang MM, Kothakota S, Haislan L, Linemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 3686; 428bp; English.
PS
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 882 AA;

Query Match 64.3%; Score 18; DB 8; Length 882;
Best Local Similarity 16.7%; Pred. No. 7.7e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
Db 763 CTTATTATTAC 774

RESULT 19
AAO14246
ID AAO14246 standard; protein; 925 AA.
XX
AC AAO14246;
XX
DT 10-MAY-2002 (first entry)
XX
DE Human presenilin enhancer protein pen-1B derived protein SEQ ID NO: 25.
XX
KW Human; fruit fly; mouse; rat; cow; presenilin enhancer protein; pen;
KW Alzheimer's disease; pen-1; pen-1B; pen-2; Aph-2; amyloid beta.
XX
OS Homo sapiens.
OS Synthetic.
OS
PN WO200185912-A2.
XX
PD 15-NOV-2001.
XX
PF 03-MAY-2001; 2001WO-US014648.
XX
PR 05-MAY-2000; 2000US-00568942.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Curtis DT, Francis GR, Ellis MC, Ruddy DA, Nicoll SM, McGrath GJ;
XX
DR WPI; 2002-062245/08.
XX

XX Presentin enhancer proteins and polynucleotides useful for modulating
PT presentin function and screening for an agent that modulates the
PI interaction of the protein to a binding target.
XX
PS Disclosure; Page 67-70; 78bp; English.
XX

XX The present invention relates to a method of detecting compounds capable
CC of altering the interaction between a presenilin enhancer protein (such
CC as pen-1, pen-1B, pen-2 and Aph-2) and presenilin. The inhibition of
CC presentin activity causes the production of amyloid beta to be reduced
CC and thus be used in the treatment of Alzheimer's disease. The present
CC sequence is a presenilin enhancer protein described in the
CC exemplification of the invention
XX
SQ Sequence 925 AA;

Query Match 64.3%; Score 18; DB 5; Length 925;
Best Local Similarity 16.7%; Pred. No. 7.8e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
Db 267 CATATTATAAC 278

RESULT 20
ADP31517
ID ADP31517 standard; protein; 1044 AA.
XX
AC ADP31517;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2284.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
OS
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 29-AUG-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471306P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486466P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
PS
PS Claim 1; SEQ ID NO 3515; 428bp; English.
XX
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
XX SQ Sequence 1044 AA;
XX
XX
XX Query Match 64.3%; Score 18; DB 8; Length 1044;
XX Best Local Similarity 16.7%; Pred. No. 8e-05;
XX Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
OY 1 CXXXXXXXXXXC 12
Db 325 CATATTAATTTC 336
XX
XX
XX RESULT 21
XX ADP31175
XX ID ADP31175 standard; protein; 1086 AA.
XX
XX ADP31175;
XX
XX 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #1942.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
XX WO2004035732-A2.
XX

PD 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406645P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411101P.
XX 17-SEP-2002; 2002US-0411111P.
XX 18-APR-2003; 2003US-0463700P.
XX 18-APR-2003; 2003US-0463708P.
XX 18-APR-2003; 2003US-0463716P.
XX 18-APR-2003; 2003US-0463732P.
XX 02-MAY-2003; 2003US-0467199P.
XX 02-MAY-2003; 2003US-0467201P.
XX 02-MAY-2003; 2003US-0467203P.
XX 02-MAY-2003; 2003US-0467230P.
XX 19-MAY-2003; 2003US-0471306P.
XX 19-MAY-2003; 2003US-0471336P.
XX 22-MAY-2003; 2003US-0472420P.
XX 22-MAY-2003; 2003US-0472430P.
XX 09-JUN-2003; 2003US-0476609P.
XX 09-JUN-2003; 2003US-0476641P.
XX 08-JUL-2003; 2003US-0485218P.
XX 08-JUL-2003; 2003US-0485223P.
XX 08-JUL-2003; 2003US-0485224P.
XX 08-JUL-2003; 2003US-0485325P.
XX 08-JUL-2003; 2003US-0485352P.
XX 14-JUL-2003; 2003US-0486466P.
XX 14-JUL-2003; 2003US-0486480P.
XX 15-JUL-2003; 2003US-0486891P.
XX 15-JUL-2003; 2003US-0486960P.
XX 08-AUG-2003; 2003US-0493341P.
XX 08-AUG-2003; 2003US-0493370P.
XX 08-AUG-2003; 2003US-0493573P.
XX 08-AUG-2003; 2003US-0493577P.
XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halembeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
XX Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3173; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 1086 AA;
Query Match 64.3%; Score 18; DB 8; Length 1086;
Best Local Similarity 16.7%; Pred. No. 8e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXXC 12
DB 189 CATTAATATTC 200
RESULT 22
ADP31508
ID ADP31508 standard; protein; 1113 AA.
XX
XX ADP31508;
AC
XX
DT 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #2275.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halembeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
XX Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 3506; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.

XX SQ Sequence 1113 AA;
Query Match 64.3%; Score 18; DB 8; Length 1113;
Best Local Similarity 16.7%; Pred. No. 8.1e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
OY 1 CXXXXXXXXXXC 12
DB 18 CATAAATTTTAAAC 29
RESULT 23
ADA15725 ADA15725 standard; protein; 1128 AA.
XX AC ADA15725;
XX DT 06-NOV-2003 (first entry)
XX DE C. elegans neuromuscular junction GABA receptor complex subunit #6.
XX NM Nematode;
XX KW neuromuscular junction gamma-aminobutyric acid receptor complex;
XX KM GABA receptor; parasitic plant pathogen; agricultural industry;
XX KM crop protection; soil treatment.
XX OS Caenorhabditis elegans.
XX PN US2003065144-A1.
XX PD 03-APR-2003.
XX PF 24-MAY-2002; 2002US-00156240.
XX PR 09-NOV-1998; 98US-0107127P.
XX PR 08-NOV-1999; 99US-00436063.
XX PA (UTAH) UNITV UTAH RES FOUND.
XX PI Bamber BA, Jorgensen EM;
XX PI WPI; 2003-540802/51.
XX DR N-PSDB; ADA15726.
XX DR New nematode neuromuscular junction GABA receptor complex, useful for
XX PT crop protection or soil treatment.
XX PS Claim 21; Page 54-56; 84pp; English.
XX CC The present invention relates to a nematode neuromuscular junction gamma-
XX CC aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant
XX CC pathogens which can cause major damage to crops in the agricultural
XX CC industry. The nematode neuromuscular junction GABA receptor complex is
XX CC useful for crop protection or soil treatment. The present sequence
XX CC represents a Caenorhabditis elegans neuromuscular junction GABA receptor
XX CC complex subunit.
XX SQ Sequence 1128 AA;
Query Match 64.3%; Score 18; DB 6; Length 1128;
Best Local Similarity 16.7%; Pred. No. 8.1e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
OY 1 CXXXXXXXXXXC 12
DB 390 CAAATATTTTTC 401
RESULT 24
ADP30741
XX ID ADP30741 standard; protein; 1134 AA.
XX

AC ADP30741;
XX 12-AUG-2004 (first entry)
XX DE Human secreted protein SEQ ID #1508.
XX NM Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX KM cancer; inflammatory; immune; human secreted protein.
XX OS Homo sapiens.
XX PN WO2004035732-A2.
XX PD 29-APR-2004.
XX PF 28-AUG-2003; 2003WO-US026780.
XX PR 29-AUG-2002; 2002US-0406576P.
XX PR 29-AUG-2002; 2002US-0406579P.
XX PR 29-AUG-2002; 2002US-0406585P.
XX PR 29-AUG-2002; 2002US-0406588P.
XX PR 29-AUG-2002; 2002US-0406608P.
XX PR 29-AUG-2002; 2002US-0406611P.
XX PR 29-AUG-2002; 2002US-0406612P.
XX PR 29-AUG-2002; 2002US-0406615P.
XX PR 29-AUG-2002; 2002US-0406640P.
XX PR 29-AUG-2002; 2002US-0406642P.
XX PR 29-AUG-2002; 2002US-0406646P.
XX PR 29-AUG-2002; 2002US-0406653P.
XX PR 29-AUG-2002; 2002US-0406655P.
XX PR 29-AUG-2002; 2002US-0406666P.
XX PR 17-SEP-2002; 2002US-0410944P.
XX PR 17-SEP-2002; 2002US-0410947P.
XX PR 17-SEP-2002; 2002US-0410948P.
XX PR 17-SEP-2002; 2002US-0410949P.
XX PR 17-SEP-2002; 2002US-0410953P.
XX PR 17-SEP-2002; 2002US-0410957P.
XX PR 17-SEP-2002; 2002US-0410958P.
XX PR 17-SEP-2002; 2002US-0410959P.
XX PR 17-SEP-2002; 2002US-0410960P.
XX PR 17-SEP-2002; 2002US-0410961P.
XX PR 17-SEP-2002; 2002US-0410962P.
XX PR 17-SEP-2002; 2002US-0410963P.
XX PR 17-SEP-2002; 2002US-0411022P.
XX PR 17-SEP-2002; 2002US-0411023P.
XX PR 17-SEP-2002; 2002US-0411024P.
XX PR 17-SEP-2002; 2002US-0411032P.
XX PR 17-SEP-2002; 2002US-0411035P.
XX PR 17-SEP-2002; 2002US-0411037P.
XX PR 17-SEP-2002; 2002US-0411041P.
XX PR 17-SEP-2002; 2002US-0411045P.
XX PR 17-SEP-2002; 2002US-0411046P.
XX PR 17-SEP-2002; 2002US-0411048P.
XX PR 17-SEP-2002; 2002US-0411052P.
XX PR 17-SEP-2002; 2002US-0411055P.
XX PR 17-SEP-2002; 2002US-0411073P.
XX PR 17-SEP-2002; 2002US-0411082P.
XX PR 17-SEP-2002; 2002US-0411101P.
XX PR 17-SEP-2002; 2002US-0411111P.
XX PR 18-APR-2003; 2003US-0463700P.
XX PR 18-APR-2003; 2003US-0463708P.
XX PR 18-APR-2003; 2003US-0463716P.
XX PR 18-APR-2003; 2003US-0463732P.
XX PR 02-MAY-2003; 2003US-0467199P.
XX PR 02-MAY-2003; 2003US-0467201P.
XX PR 02-MAY-2003; 2003US-0467203P.
XX PR 02-MAY-2003; 2003US-0467206P.
XX PR 19-MAY-2003; 2003US-0471306P.
XX PR 19-MAY-2003; 2003US-0471336P.
XX PR 22-MAY-2003; 2003US-0472420P.
XX PR 22-MAY-2003; 2003US-0472430P.
XX PR 09-JUN-2003; 2003US-0476609P.
XX PR 09-JUN-2003; 2003US-0476641P.

[illegible]

PR	17-SEP-2002;	2002US-0411023P.
PR	17-SEP-2002;	2002US-0411024P.
PR	17-SEP-2002;	2002US-0411032P.
PR	17-SEP-2002;	2002US-0411035P.
PR	17-SEP-2002;	2002US-0411037P.
PR	17-SEP-2002;	2002US-0411041P.
PR	17-SEP-2002;	2002US-0411045P.
PR	17-SEP-2002;	2002US-0411046P.
PR	17-SEP-2002;	2002US-0411048P.
PR	17-SEP-2002;	2002US-0411052P.
PR	17-SEP-2002;	2002US-0411055P.
PR	17-SEP-2002;	2002US-0411073P.
PR	17-SEP-2002;	2002US-0411082P.
PR	17-SEP-2002;	2002US-0411011P.
PR	17-SEP-2002;	2002US-0411111P.
PR	18-APR-2003;	2003US-0463700P.
PR	18-APR-2003;	2003US-0463708P.
PR	18-APR-2003;	2003US-0463716P.
PR	18-APR-2003;	2003US-0463732P.
PR	02-MAY-2003;	2003US-0467199P.
PR	02-MAY-2003;	2003US-0467201P.
PR	02-MAY-2003;	2003US-0467203P.
PR	02-MAY-2003;	2003US-0467230P.
PR	19-MAY-2003;	2003US-0471306P.
PR	19-MAY-2003;	2003US-0471336P.
PR	22-MAY-2003;	2003US-0472420P.
PR	22-MAY-2003;	2003US-0472430P.
PR	09-JUN-2003;	2003US-0476609P.
PR	09-JUN-2003;	2003US-0476641P.
PR	08-JUL-2003;	2003US-0485218P.
PR	08-JUL-2003;	2003US-0485223P.
PR	08-JUL-2003;	2003US-0485224P.
PR	08-JUL-2003;	2003US-0485325P.
PR	14-JUL-2003;	2003US-0486446P.
PR	14-JUL-2003;	2003US-0486480P.
PR	15-JUL-2003;	2003US-0486891P.
PR	15-JUL-2003;	2003US-0486960P.
PR	08-AUG-2003;	2003US-0493341P.
PR	08-AUG-2003;	2003US-0493370P.
PR	08-AUG-2003;	2003US-0493579P.
PR	08-AUG-2003;	2003US-0493577P.
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
XX		
PI	Williams LT, Chu K, Lee E, Heetir K, Beaurang PA, Behrens D,	
PI	Halenbeck RF, Huang MW, Kotnakota S, Hsiehan L, Linnemann T;	
PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;	
XX		
DR	WPI, 2004-348438/32.	
XX		
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	genetic, bacterial and viral diseases.	
XX		
PS	Claim 1; SEQ ID NO 3044; 428bp; English.	
CC		
CC	The present invention relates to an isolated nucleic acid molecule	
CC	encoding a polypeptide which is believed to be cytosolic,	
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	composition and methods are useful for diagnosing, preventing and	
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,	
CC	immune, metabolic, genetic, bacterial and viral diseases. The present	
CC	sequence represents a human secreted protein. The present sequence is	
CC	available on WIPWEB and is not in the specification.	
XX		
SQ	Sequence 1168 AA;	
XX		
Query Match	64.3%;	Score 18; DB 8; Length 1168;
Best Local Similarity	16.7%;	Pred. No. 8.2e-05;
Matches	2; Conservative	0; Mismatches 10; Indels 0; Gaps 0
QY	1 CXXXXXXXXXXC 12	

DB 223 CAATTAATTAC 234

RESULT 27

ADP30993

ID ADP30993 standard; protein; 1191 AA.

XX

AC ADP30993;

XX

DT 12-AUG-2004 (first entry)

XX

DE Human secreted protein SEQ ID #1760.

XX

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

XX

OS cancer; inflammatory; immune; human secreted protein.

XX

OS Homo sapiens.

XX

PN WO2004035732-A2.

XX

PD 29-APR-2004.

XX

PF 28-AUG-2003; 2003WO-US026780.

XX

PR 29-AUG-2002; 2002US-0406576P.

XX

PR 29-AUG-2002; 2002US-0406579P.

XX

PR 29-AUG-2002; 2002US-0406585P.

XX

PR 29-AUG-2002; 2002US-0406588P.

XX

PR 29-AUG-2002; 2002US-0406608P.

XX

PR 29-AUG-2002; 2002US-0406611P.

XX

PR 29-AUG-2002; 2002US-0406612P.

XX

PR 29-AUG-2002; 2002US-0406616P.

XX

PR 29-AUG-2002; 2002US-0406640P.

XX

PR 29-AUG-2002; 2002US-0406642P.

XX

PR 29-AUG-2002; 2002US-0406646P.

XX

PR 29-AUG-2002; 2002US-0406653P.

XX

PR 29-AUG-2002; 2002US-0406655P.

XX

PR 29-AUG-2002; 2002US-0406666P.

XX

PR 17-SEP-2002; 2002US-0410946P.

XX

PR 17-SEP-2002; 2002US-0410947P.

XX

PR 17-SEP-2002; 2002US-0410948P.

XX

PR 17-SEP-2002; 2002US-0410949P.

XX

PR 17-SEP-2002; 2002US-0410953P.

XX

PR 17-SEP-2002; 2002US-0410957P.

XX

PR 17-SEP-2002; 2002US-0410958P.

XX

PR 17-SEP-2002; 2002US-0410959P.

XX

PR 17-SEP-2002; 2002US-0410960P.

XX

PR 17-SEP-2002; 2002US-0410961P.

XX

PR 17-SEP-2002; 2002US-0410962P.

XX

PR 17-SEP-2002; 2002US-0411019P.

XX

PR 17-SEP-2002; 2002US-0411022P.

XX

PR 17-SEP-2002; 2002US-0411023P.

XX

PR 17-SEP-2002; 2002US-0411024P.

XX

PR 17-SEP-2002; 2002US-0411025P.

XX

PR 17-SEP-2002; 2002US-0411032P.

XX

PR 17-SEP-2002; 2002US-0411035P.

XX

PR 17-SEP-2002; 2002US-0411037P.

XX

PR 17-SEP-2002; 2002US-0411041P.

XX

PR 17-SEP-2002; 2002US-0411045P.

XX

PR 17-SEP-2002; 2002US-0411046P.

XX

PR 17-SEP-2002; 2002US-0411048P.

XX

PR 17-SEP-2002; 2002US-0411052P.

XX

PR 17-SEP-2002; 2002US-0411055P.

XX

PR 17-SEP-2002; 2002US-0411073P.

XX

PR 17-SEP-2002; 2002US-0411073P.

XX

PR 17-SEP-2002; 2002US-0411082P.

XX

PR 17-SEP-2002; 2002US-0411101P.

XX

PR 17-SEP-2002; 2002US-0411111P.

XX

PR 18-APR-2003; 2003US-0463700P.

XX

PR 18-APR-2003; 2003US-0463708P.

XX

PR 18-APR-2003; 2003US-0463716P.

XX

PR 18-APR-2003; 2003US-0463732P.

XX

PR 02-MAY-2003; 2003US-0467199P.

XX

PR 02-MAY-2003; 2003US-0467201P.

XX

PR 02-MAY-2003; 2003US-0467203P.

XX

PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.

PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476409P.

PR 09-JUN-2003; 2003US-0476412P.

PR 08-JUL-2003; 2003US-0485218P.

PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485224P.

PR 08-JUL-2003; 2003US-0485225P.

PR 14-JUL-2003; 2003US-0486446P.

PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486960P.

PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.

PR 08-AUG-2003; 2003US-0493577P.

XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX

PI Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D,

PI Halenbeck RF, Huang MM, Kotnakota S, Haisan L, Lannemann T,

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

XX

DR WPI; 2004-348438/32.

XX

PT New nucleic acid molecule for diagnosing, preventing or treating diseases

PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,

PT genetic, bacterial and viral diseases.

XX

PS Claim 1; SEQ ID NO 2991; 428bp; English.

XX

CC The present invention relates to an isolated nucleic acid molecule

CC encoding a polypeptide which is believed to be cytostatic,

CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and

CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC immune, metabolic, genetic, bacterial and viral diseases. The present

CC sequence represents a human secreted protein. The present sequence is

CC available on WIPWEB and is not in the specification.

XX

SQ Sequence 1191 AA;

XX

Query Match 64.3%; Score 18; DB 8; Length 1191;

Best Local Similarity 16.7%; Pred. No. 8.2e-05;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 12

DB 821 CTTATTAATTATC 832

RESULT 28

ADP31533

ID ADP31533 standard; protein; 1260 AA.

XX

AC ADP31533;

XX

DT 12-AUG-2004 (first entry)

XX

DE Human secreted protein SEQ ID #2300.

XX

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

XX

OS cancer; inflammatory; immune; human secreted protein.

XX

OS Homo sapiens.

XX

PN WO2004035732-A2.

XX

PD 29-APR-2004.

XX

PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471366P.
PR 22-MAY-2003; 2003US-0472430P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476611P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485252P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Hsieh L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
DR
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3531; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 1260 AA;
XX
Query Match 64.3%; Score 18; DB 8; Length 1260;
Best Local Similarity 16.7%; Pred. No. 8.3e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXXC 12
Db 757 CAATTAATAATAC 768
RESULT 29
ADP30675
ID ADP30675 standard; protein; 1289 AA.
XX
AC ADP30675;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1442.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411053P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411075P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-047136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485252P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-048691P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

XX
PA (FIVE-) FIVE PRIME.THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Hallenbeck RF, Huang MM, Kotchikora S, Haishan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 2673; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.

Sequence 1289 AA;

```

Query Match      64.3%  Score 18: DB 8; Length 1289;
Best Local Similarity 16.7%  Pred. No. 8.3e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0.

QY      1 CXXXXXXXXXXC 12
Db      1076 CATATATAATTC 1087

RESULT 30
ADP31357
ID      ADP31357 standard; protein; 1437 AA.
XX
AC      ADP31357;
XX
DT      12-AUG-2004 (first entry)
XX
DE      Human secreted protein SEQ ID #2124.
XX
KW      Cytosolic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW      cancer; inflammatory; immune; human secreted protein.
XX
OS      Homo sapiens.
XX
PN      WO2004035732-A2.
XX
PD      29-APR-2004.
XX
PF      28-AUG-2003; 2003WO-US026780.
XX
PR      29-AUG-2002; 2002US-0406576P.
PR      29-AUG-2002; 2002US-0406579P.
PR      29-AUG-2002; 2002US-0406585P.
PR      29-AUG-2002; 2002US-0406588P.
PR      29-AUG-2002; 2002US-0406608P.
PR      29-AUG-2002; 2002US-0406611P.
PR      29-AUG-2002; 2002US-0406612P.
PR      29-AUG-2002; 2002US-0406616P.
PR      29-AUG-2002; 2002US-0406640P.
PR      29-AUG-2002; 2002US-0406642P.
PR      29-AUG-2002; 2002US-0406646P.
PR      29-AUG-2002; 2002US-0406653P.
PR      29-AUG-2002; 2002US-0406655P.
PR      29-AUG-2002; 2002US-0406666P.
PR      17-SEP-2002; 2002US-0410946P.
PR      17-SEP-2002; 2002US-0410947P.
PR      17-SEP-2002; 2002US-0410948P.
PR      17-SEP-2002; 2002US-0410949P.
PR      17-SEP-2002; 2002US-0410953P.
PR      17-SEP-2002; 2002US-0410957P.
PR      17-SEP-2002; 2002US-0410958P.
PR      17-SEP-2002; 2002US-0410959P.
PR      17-SEP-2002; 2002US-0410960P.
PR      17-SEP-2002; 2002US-0410961P.
PR      17-SEP-2002; 2002US-0410962P.
PR      17-SEP-2002; 2002US-0411022P.
PR      17-SEP-2002; 2002US-0411024P.
PR      17-SEP-2002; 2002US-0411035P.
PR      17-SEP-2002; 2002US-0411037P.
PR      17-SEP-2002; 2002US-0411041P.
PR      17-SEP-2002; 2002US-0411045P.
PR      17-SEP-2002; 2002US-0411046P.
PR      17-SEP-2002; 2002US-0411048P.
PR      17-SEP-2002; 2002US-0411052P.
PR      17-SEP-2002; 2002US-0411055P.
PR      17-SEP-2002; 2002US-0411073P.
PR      17-SEP-2002; 2002US-0411082P.
PR      17-SEP-2002; 2002US-0411101P.
PR      17-SEP-2002; 2002US-0411111P.

```


PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3175; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.
XX
SQ Sequence 1454 AA;
XX
Query Match 64.3%; Score 18; DB 8; Length 1454;
Best Local Similarity 16.7%; Pred. No. 8.5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXXC 12
DB 1094 CATTATATATAC 1105
XX
RESULT 32
ADP30557
ID ADP30557 standard; protein; 1480 AA.
XX
AC ADP30557;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1324.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471366P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476618P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485252P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486910P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 2555; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.

XX
SQ Sequence 1480 AA;

Query Match 64.3%; Score 18; DB 8; Length 1480;
Best Local Similarity 16.7%; Pred. No. 8.6e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 12

Db 1286 CTPAAATATTAAC 1297

RESULT 33

ABBO9437 standard; protein; 1588 AA.

XX
AC ABBO9437;

DT 01-JUL-2002 (first entry)

DE H. influenzae DXR related polypeptide sequence.

XX DXR; reductoisomerase; enzyme; non-mevalonate isoprenoid; menaquinone;
XX ubiquinone; vitamin; ear infection; conjunctivitis; meningitis;
KW pneumonia; conjunctivitis; bacteraemia; sinusitis; pleural empyema;
KW endocarditis; epiglottitis.

XX
OS Haemophilus influenzae.

XX
FH Key Location/Qualifiers

FT Region 241..1431
FT /note="region that appears to be accidentally inserted
FT into the sequence, consisting the DXR encoding DNA
FT sequence represented as an amino acid sequence in three
FT letter code"

PD WO200211673-A2.

XX 14-FEB-2002.

XX 09-AUG-2001; 2001WO-US024950.

XX 09-AUG-2000; 2000US-022309P.

PA (SMIT) SMITHKLINE BEECHAM CORP.

PA (SMITK) SMITHKLINE BEECHAM PLC.

PI Jaworski DD, Payne DJ, Slater-Radoski CE, Yan K;

DR WPI; 2002-241698/29.

PT Modulating Haemophilus influenzae DXR reductoisomerase enzyme activity,
PT useful for treating mammals or tissues infected with H. influenzae (e.g.
PT ear infections or pneumonia) by contacting the enzyme with a modulator of
PT its activity.

PS Disclosure; Page 40-44; 44pp; English.

XX The invention relates to modulating an activity of a DXR reductoisomerase
CC enzyme of Haemophilus influenzae, comprising contacting the enzyme with a
CC compound that modulates non-mevalonate isoprenoid biosynthesis -
CC synthesis of menaquinone or ubiquinone. Compounds of the invention act as
CC virucides. The method is useful for treating a mammal or mammalian tissue
CC infected with H. influenzae having DXR reductoisomerase enzyme, e.g. a
CC human or a domestic animal. In particular, the method is useful for
CC treating ear infections, conjunctivitis, meningitis, pneumonia,
CC conjunctivitis, bacteraemia, sinusitis, pleural empyema, endocarditis and
CC epiglottitis. The current sequence represents a H. influenzae DXR

CC reductoisomerase enzyme related polypeptide sequence. Note: The current
CC sequence contains within it the amino acid sequence given in record
CC ABBO9436 (DXR enzyme), but this is broken up by a large insertion that
CC appears to be accidentally inserted into the sequence, consisting the DXR
CC encoding DNA sequence represented as an amino acid sequence in three
CC letter code

XX
SQ Sequence 1588 AA;

Query Match 64.3%; Score 18; DB 5; Length 1588;
Best Local Similarity 16.7%; Pred. No. 8.7e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 12

Db 431 CTPAAATTTTAC 442

RESULT 34

ADP30660 standard; protein; 1617 AA.

XX
AC ADP30660;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #1427.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.

XX
OS Homo sapiens.

XX
PN WO2004035732-A2.

XX 29-APR-2004.

PF 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406640P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406646P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

XX 29-AUG-2002; 2002US-0410946P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410957P.

XX 17-SEP-2002; 2002US-0410959P.

XX 17-SEP-2002; 2002US-0410960P.

XX 17-SEP-2002; 2002US-0410961P.

XX 17-SEP-2002; 2002US-0410962P.

XX 17-SEP-2002; 2002US-0411019P.

XX 17-SEP-2002; 2002US-0411022P.

XX 17-SEP-2002; 2002US-0411023P.

XX 17-SEP-2002; 2002US-0411024P.

XX 17-SEP-2002; 2002US-0411032P.

XX 17-SEP-2002; 2002US-0411035P.

XX 17-SEP-2002; 2002US-0411037P.

XX 17-SEP-2002; 2002US-0411041P.

XX 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.
 PR 17-SEP-2002; 2002US-0411048P.
 PR 17-SEP-2002; 2002US-0411052P.
 PR 17-SEP-2002; 2002US-0411055P.
 PR 17-SEP-2002; 2002US-0411073P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411101P.
 PR 17-SEP-2002; 2002US-0411111P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 02-MAY-2003; 2003US-0467230P.
 PR 19-MAY-2003; 2003US-0471366P.
 PR 19-MAY-2003; 2003US-0471366P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 22-MAY-2003; 2003US-0472430P.
 PR 09-JUN-2003; 2003US-0476609P.
 PR 09-JUN-2003; 2003US-0476641P.
 PR 08-JUL-2003; 2003US-0485218P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485224P.
 PR 08-JUL-2003; 2003US-0485325P.
 PR 14-JUL-2003; 2003US-0486446P.
 PR 14-JUL-2003; 2003US-0486480P.
 PR 15-JUL-2003; 2003US-0486891P.
 PR 15-JUL-2003; 2003US-0486960P.
 PR 08-AUG-2003; 2003US-0493341P.
 PR 08-AUG-2003; 2003US-0493370P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
 PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

DR WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 2658; 428bp; English.

CC The present invention relates to an isolated nucleic acid molecule
 CC encoding a polypeptide which is believed to be cytostatic,
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is
 CC available on WIPOMB and is not in the specification.

XX Sequence 1617 AA;

QY Query Match 64.3%; Score 18; DB 8; Length 1617;
 Best Local Similarity 16.7%; Pred. No. 8.7e-05;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
 Db 71 CTATATTAATTC 82

RESULT 35

ADA15715
 ID ADA15715 standard; protein; 1652 AA.
 XX

AC ADA15715;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 XX C. elegans neuromuscular junction GABA receptor complex subunit #1.
 DE
 XX
 XX Nematode;
 KW neuromuscular junction gamma-aminobutyric acid receptor complex;
 KW GABA receptor; parasitic plant pathogen; agricultural industry;
 KW crop protection; soil treatment.
 OS
 OS Caenorhabditis elegans.
 PN
 XX US2003065144-A1.
 PD
 PD 03-APR-2003.
 PF
 PF 24-MAY-2002; 2002US-00156240.
 XX
 XX 09-NOV-1998; 98US-0107727P.
 PR
 PR 08-NOV-1999; 99US-00436063.
 XX
 PA (UTAH) UNIV UTAH RES FOUND.
 XX
 XX Bamber BA, Jorgensen EM;
 PI
 XX WPI; 2003-540802/51.
 DR
 XX
 XX New nematode neuromuscular junction GABA receptor complex, useful for
 PT crop protection or soil treatment.
 PT
 PT Claim 21; Page 20-24; 84p; English.

CC The present invention relates to a nematode neuromuscular junction gamma-
 CC aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant
 CC pathogens which can cause major damage to crops in the agricultural
 CC industry. The nematode neuromuscular junction GABA receptor complex is
 CC useful for crop protection or soil treatment. The present sequence
 CC represents a Caenorhabditis elegans neuromuscular junction GABA receptor
 CC complex subunit.

XX Sequence 1652 AA;

QY Query Match 64.3%; Score 18; DB 6; Length 1652;
 Best Local Similarity 16.7%; Pred. No. 8.8e-05;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
 Db 891 CAATATTTTTC 902

RESULT 36

ADP30654
 ID ADP30654 standard; protein; 1725 AA.

XX ADP30654;

XX 12-AUG-2004 (first entry)

XX Human secreted protein SEQ ID #1421.

DE Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
 KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

PN WO2004035732-A2.

PD 29-APR-2004.

PF 28-AUG-2003; 2003WO-US026780.

PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471106P.
PR 19-MAY-2003; 2003US-0471136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486466P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486919P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T,

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2652; 428pp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 1725 AA;
XX
Query Match 64.3%; Score 18; DB 8; Length 1725;
Best Local Similarity 16.7%; Pred. No. 8.9e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXXC 12
DB 828 CTATTTTAAAC 839
RESULT 37
ADP30642
ID ADP30642 standard, protein; 1833 AA.
XX
XX ADP30642;
XX
DT 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #1409.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
XX WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 26-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485233P.
PR 08-JUL-2003; 2003US-0485244P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Halaban L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2640; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cytostatic,
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPOMEB and is not in the specification.
XX
XX Sequence 1833 AA;
SQ

Query Match 64.3%; Score 18; DB 8; Length 1833;

Best Local Similarity 16.7%; Pred. No. 9e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXX 12
Db 462 CTATTTTAAAC 473

RESULT 38
ADP31178
ID ADP31178 standard; protein; 2088 AA.

XX ADP31178;
XX

DT 12-AUG-2004 (first entry)
XX

DE Human secreted protein SEQ ID #1945.
XX

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.

XX WO2004035732-A2.
XX

PD 29-APR-2004.
XX

PF 28-AUG-2003; 2003WO-US026780.
XX

XX 29-AUG-2002; 2002US-0406576P.
XX

XX 29-AUG-2002; 2002US-0406579P.
XX

XX 29-AUG-2002; 2002US-0406585P.
XX

XX 29-AUG-2002; 2002US-0406588P.
XX

XX 29-AUG-2002; 2002US-0406608P.
XX

XX 29-AUG-2002; 2002US-0406611P.
XX

XX 29-AUG-2002; 2002US-0406612P.
XX

XX 29-AUG-2002; 2002US-0406616P.
XX

XX 29-AUG-2002; 2002US-0406642P.
XX

XX 29-AUG-2002; 2002US-0406646P.
XX

XX 29-AUG-2002; 2002US-0406653P.
XX

XX 29-AUG-2002; 2002US-0406655P.
XX

XX 29-AUG-2002; 2002US-0406666P.
XX

XX 17-SEP-2002; 2002US-0410946P.
XX

XX 17-SEP-2002; 2002US-0410947P.
XX

XX 17-SEP-2002; 2002US-0410948P.
XX

XX 17-SEP-2002; 2002US-0410949P.
XX

XX 17-SEP-2002; 2002US-0410953P.
XX

XX 17-SEP-2002; 2002US-0410958P.
XX

XX 17-SEP-2002; 2002US-0410959P.
XX

XX 17-SEP-2002; 2002US-0410960P.
XX

XX 17-SEP-2002; 2002US-0410961P.
XX

XX 17-SEP-2002; 2002US-0410962P.
XX

XX 17-SEP-2002; 2002US-0410966P.
XX

PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485252P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
PR New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3176; 428pp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 2088 AA;
Query Match 64.3%; Score 18; DB 8; Length 2088;
Best Local Similarity 16.7%; Pred. No. 9.2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXXC 12
Db 1094 CATTATATATAC 1105

RESULT 39
ADP31327
ID ADP31327 standard; protein; 2127 AA.
XX
XX ADP31327;
AC
XX
XX 12-AUG-2004 (first entry)
DT
XX
XX Human secreted protein SEQ ID #2094.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
OS

XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411015P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 02-MAY-2003; 2003US-0467236P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
DR
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PR genetic, bacterial and viral diseases.
XX
XX
PS Claim 1; SEQ ID NO 3325; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 2127 AA;
XX
XX
XX Query Match 64.3%; Score 18; DB 8; Length 2127;
XX Best Local Similarity 16.7%; Pred. No. 9.2e-05;
XX Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
QY 1 CXXXXXXXXXXC 12
XX
DB 1464 CTTATTAAAAAC 1475
XX
XX
XX RESULT 40
XX ADP66690
XX ID ADP66690 standard; protein; 2484 AA.
XX
XX AC ADP66690;
XX
XX
XX 26-AUG-2004 (first entry)
XX
XX
XX Human mismatch repair protein MLH1.
XX
XX
XX PMS1; PMS2; PMSR3; PMSR2; PMSR6; MLH1; GTBP; MSH3; MSH1; PMSR;
XX immunoglobulin; mismatch repair protein; human.
XX
XX
XX Homo sapiens.
XX
XX
XX WO2004046330-A2.
XX
XX
XX 03-JUN-2004.
XX
XX
XX 14-NOV-2003; 2003WO-US036702.
XX
XX
XX 15-NOV-2002; 2002US-0427165P.
XX
XX 10-SEP-2003; 2003US-0501650P.
XX
XX
XX (MORP-) MORPHOTEK INC.
XX
XX
XX Graessl L, Liang S, Nicolaides NE, Saas PM;
XX WPI; 2004-440979/41.
XX
XX
XX Producing mammalian expression cells producing high-affinity antibodies
PT from immunized immunoglobulin-producing cells, by combining cells with
PT antigen, forming parental hybridoma cells and hypermutated hybridoma
PT cells, using myeloma cells.
XX
XX
XX Claim 135; SEQ ID NO 12; 213bp; English.

XX
XX The invention relates to producing mammalian expression cells e.g.,
CC hybridoma cells producing high-affinity and high titer antibodies from in
CC vitro immunized immunoglobulin-producing cells. The method involves:
CC combining donor cells comprising immunoglobulin (Ig)-producing cells with
CC an immunogenic antigen in vitro, fusing the Ig-producing cells with
CC myeloma cells to form parental hybridoma cells, where the hybridoma cells
CC express a dominant negative allele of a mismatch repair gene, incubating
CC the parental hybridoma cells to allow for mutagenesis, thus forming
CC hypermutated hybridoma cells, performing a screen for binding of
CC antibodies to antigen for antibodies produced from the hypermutated
CC hybridoma cells, and selecting hypermutated hybridoma cells that produce
CC antibodies with greater affinity for the antigen than antibodies produced
CC by the parental hybridoma cells, thus producing hybridoma cells producing
CC high-affinity antibodies. In the method, the dominant negative allele of
CC a mismatch repair gene comprises a dominant negative allele of a gene
CC chosen from PMS2, PMS1, PMSR3, PMSR2, PMSR6, MLH1, GTBP, MSH3, or
CC MSH1, and homologues of PMSR genes. The method is useful for producing
CC mammalian expression cells e.g., hybridoma cells producing high-affinity
CC and high titer antibodies from in vitro immunized immunoglobulin-
CC producing cells. The present sequence represents a human mismatch repair
CC protein MLH1.
XX
XX
SQ Sequence 2484 AA;
XX
XX
XX Query Match 64.3%; Score 18; DB 8; Length 2484;
XX Best Local Similarity 16.7%; Pred. No. 9.6e-05;
XX Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
QY 1 CXXXXXXXXXXC 12
XX
DB 515 CTTTAAAAAATC 526
XX
XX
XX RESULT 41
XX ADA15721
XX ID ADA15721 standard; protein; 2508 AA.
XX
XX AC ADA15721;
XX
XX
XX 06-NOV-2003 (first entry)
XX
XX
XX C. elegans neuromuscular junction GABA receptor complex subunit #4.
XX
XX
XX Nematoide;
XX neuromuscular junction gamma-aminobutyric acid receptor complex;
XX GABA receptor; parasitic plant pathogen; agricultural industry;
XX crop protection; soil treatment.
XX
XX
XX Caenorhabditis elegans.
XX
XX
XX US2003065144-A1.
XX
XX
XX 03-APR-2003.
XX
XX
XX 24-MAY-2002; 2002US-00156240.
XX
XX
XX 09-NOV-1998; 98US-0107727P.
XX
XX 08-NOV-1999; 99US-00436063.
XX
XX
XX (UTAH) UNIV UTAH RES FOUND.
XX
XX
XX Bamber BA, Jorgensen EM;
XX WPI; 2003-540802/51.
XX
XX
XX N-PSDB; ADA15722.
XX
XX
XX New nematode neuromuscular junction GABA receptor complex, useful for
PT crop protection or soil treatment.
XX
XX
XX Claim 21; Page 38-44; 84bp; English.
XX
XX
XX The present invention relates to a nematode neuromuscular junction gamma-

CC aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant
CC pathogens which can cause major damage to crops in the agricultural
CC industry. The nematode neuromuscular junction GABA receptor complex is
CC useful for crop protection or soil treatment. The present sequence
CC represents a Caenorhabditis elegans neuromuscular junction GABA receptor
CC complex subunit.

SQ Sequence 2508 AA;

Query Match 64.3%; Score 18; DB 6; Length 2508;

Best Local Similarity 16.7%; Pred. No. 9.6e-05; Mismatches 10; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12

Db 1775 CAAATATTTTC 1786

RESULT 42

ADA15717
ID ADA15717 standard; protein; 2544 AA.

XX ADA15717;

DT 06-NOV-2003 (first entry)

DE C. elegans neuromuscular junction GABA receptor complex subunit #2.

XX Nematode;

KW neuromuscular junction gamma-aminobutyric acid receptor complex;

KW GABA receptor; parasitic plant pathogen; agricultural industry;

XX crop protection; soil treatment.

OS Caenorhabditis elegans.

PN US2003065144-A1.

XX 03-APR-2003.

PF 24-MAY-2002; 2002US-00156240.

PR 09-NOV-1998; 98US-0107727P.

PR 08-NOV-1999; 99US-00436063.

XX (UTAH) UNIV UTAH RES FOUND.

PI Bamber BA, Jorgensen EM;

XX WPI, 2003-540802/51.

DR N-PSDB; ADA15718.

PT New nematode neuromuscular junction GABA receptor complex, useful for

PT crop protection or soil treatment.

PS Claim 21; Page 25-31; 84pp; English.

XX The present invention relates to a nematode neuromuscular junction gamma-
CC aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant
CC pathogens which can cause major damage to crops in the agricultural
CC industry. The nematode neuromuscular junction GABA receptor complex is
CC useful for crop protection or soil treatment. The present sequence
CC represents a Caenorhabditis elegans neuromuscular junction GABA receptor
CC complex subunit.

SQ Sequence 2544 AA;

Query Match 64.3%; Score 18; DB 6; Length 2544;

Best Local Similarity 16.7%; Pred. No. 9.6e-05; Mismatches 10; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12

Db 1811 CAAATATTTTC 1822

RESULT 43

ADA15723
ID ADA15723 standard; protein; 2601 AA.

XX ADA15723;

DT 06-NOV-2003 (first entry)

DE C. elegans neuromuscular junction GABA receptor complex subunit #5.

XX Nematode;

KW neuromuscular junction gamma-aminobutyric acid receptor complex;

KW GABA receptor; parasitic plant pathogen; agricultural industry;

XX crop protection; soil treatment.

OS Caenorhabditis elegans.

PN US2003065144-A1.

XX 03-APR-2003.

PF 24-MAY-2002; 2002US-00156240.

PR 09-NOV-1998; 98US-0107727P.

PR 08-NOV-1999; 99US-00436063.

XX (UTAH) UNIV UTAH RES FOUND.

PI Bamber BA, Jorgensen EM;

XX WPI, 2003-540802/51.

DR N-PSDB; ADA15724.

PT New nematode neuromuscular junction GABA receptor complex, useful for

PT crop protection or soil treatment.

PS Claim 21; Page 46-52; 84pp; English.

XX The present invention relates to a nematode neuromuscular junction gamma-
CC aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant
CC pathogens which can cause major damage to crops in the agricultural
CC industry. The nematode neuromuscular junction GABA receptor complex is
CC useful for crop protection or soil treatment. The present sequence
CC represents a Caenorhabditis elegans neuromuscular junction GABA receptor
CC complex subunit.

SQ Sequence 2601 AA;

Query Match 64.3%; Score 18; DB 6; Length 2601;

Best Local Similarity 16.7%; Pred. No. 9.6e-05; Mismatches 10; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12

Db 1251 CAATAATATTC 1262

RESULT 44

ADP31299
ID ADP31299 standard; protein; 2833 AA.

XX ADP31299;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2066.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

XX cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

```

XX WO2004035732-A2.
PN
XX
XX
PD 29-APR-2004.
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471366P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.

```

```

PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang PM, Kothakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGB, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1, SEQ ID NO 3297; 428bp; English.
XX
XX CC The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cytostatic,
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPOMER and is not in the specification.
XX
SQ Sequence 2833 AA;
Query Match 64.3%; Score 18; DB 8; Length 2833;
Best Local Similarity 16.7%; Pred. No. 9.8e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXXC 12
DB 1876 CATTAATTATATC 1887
RESULT 45
ADP30572
ID ADP30572 standard; protein; 2835 AA.
XX
XX ADP30572;
AC
XX
XX 12-AUG-2004 (first entry)
DT
XX
DE Human secreted protein SEQ ID #1339.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.

```

PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411025P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411011P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485252P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halsebeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX MPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
PS Claim 1: SEQ ID NO 2570; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 2835 AA;
Query Match 64.3%; Score 18; DB 8; Length 2835;
Best Local Similarity 16.7%; Pred. No. 9.8e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXXC 12
Db 884 CAAAAATTATTC 895
RESULT 46
ADP30667
ID ADP30667 standard; protein; 3411 AA.
XX
AC ADP30667;
XX
XX 12-AUG-2004 (first entry)
DT XX
XX
DE Human secreted protein SEQ ID #1434.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471366P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halembeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.

PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

PS Claim 1: SEQ ID NO 2665; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOMB and is not in the specification.

SO Sequence 3411 AA;

Query Match 64.3%; Score 18; DB 8; Length 3411;
Best Local Similarity 16.7%; Pred. No. 0.0001;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12
Db 2826 CAATTATTTC 2837

RESULT 47
ABU88256
ID ABU88256 standard; protein; 4440 AA.

XX ABU88256;
XX
DT 07-JUL-2003 (first entry)

XX Novel human secreted and transmembrane PRO polypeptide #4.
DE Human; secreted and transmembrane protein: PRO; gene therapy;
XX tumour necrosis factor-alpha release; TNF-alpha release;
KW chondrocyte proliferation; chondrocyte differentiation; tumour;
KW adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS Homo sapiens.
XX
PN US2003032127-A1.
XX
PD 13-FEB-2003.
XX
PF 26-JUN-2002; 2002US-00183012.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 13-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081836P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 26-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.

PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088555P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088611P.
PR 10-JUN-1998; 98US-0088624P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089588P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095988P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-009691P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.

PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098832P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98MO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101923P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 30-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102587P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.

Query Match 64.3%; Score 18; DB 6; Length 4440;
Best Local Similarity 16.7%; Pred. No. 0.00011;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 12
Db 3858 CATTTTAATAC 3869

RESULT 48
ABU90135 standard; protein; 4440 AA.
ID ABU90135;
XX AC ABU90135;
XX DT 11-AUG-2003 (first entry)
XX DE Novel human secreted and transmembrane PRO protein #4.
XX KW Human; gene therapy; tissue typing; tumour; chondrocyte proliferation;
XX chondrocyte differentiation; tumour necrosis factor-alpha release;
KW affinity purification.

OS Homo sapiens.
XX
PN US2003036147-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-00187741.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-008579P.
PR 15-MAY-1998; 98US-008580P.
PR 15-MAY-1998; 98US-0085882P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.

PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 09-JUN-1998; 98US-0088217P.
PR 10-JUN-1998; 98US-008855P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-009029P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-0010541P.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096575P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.

PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0098602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101933P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 23-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102230P.
PR 29-SEP-1998; 98US-0102311P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103497P.
PR 07-OCT-1998; 98US-00168978.

Query Match 64.3%; Score 18; DB 6; Length 4440;
Best Local Similarity 16.7%; Pred. No. 0.00011;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 12
DB 3858 CATTTTTATATAC 3869

RESULT 49
ABU96437
ID ABU96437 strand: protein; 4440 AA.

AC ABU96437;
DT 25-JUL-2003 (first entry)

DE Novel human secreted and transmembrane protein #4.

XX Human; secreted and transmembrane protein; PRO; transgenic animal;
KW knockout; chromosome identification; tissue typing; tumour;
KW chondrocyte proliferation; chondrocyte differentiation;
KW tumor necrosis factor-alpha release stimulator.

XX Homo sapiens.

XX US2003036144-A1.

XX 20-FEB-2003.

XX

PF 01-JUL-2002; 2002US-00187601.
XX 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081070P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 05-MAY-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 07-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.


```
PR 09-JUN-1998; 98US-008655P.
PR 10-JUN-1998; 98US-008722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091526P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 04-AUG-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095989P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 26-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.

PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-009988P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 23-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 64.3%; Score 18; DB 6; Length 4440;
Best Local Similarity 16.7%; Pred. No. 0.00011;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
Db 3858 CATTTTAATAC 3869

RESULT 50
ABU99046
ID ABU99046 standard; protein; 4440 AA.
XX
AC ABU99046;
XX
DT 01-AUG-2003 (first entry)
XX
DE
DE Novel human secreted and transmembrane protein #4.
XX
XX Human; secreted and transmembrane protein; PRO; cytosolic; gene therapy;
KW chondrocyte stimulator; tumour; adrenal tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; TNF-alpha release;
KW tumour necrosis factor alpha release; chondrocyte cell proliferation;
KW chondrocyte cell differentiation; pharmaceutical; diagnostic; biosensor;
KW bioreactor.
XX
XX Homo sapiens.
OS
XX
XX
XX PN US2003013153-A1.
XX
XX PD 16-JAN-2003.
XX
XX PF 19-JUN-2002; 2002US-00175737.
XX
XX 18-SEP-1997; 97US-0059263P.
XX
XX 18-SEP-1997; 97US-0059266P.
```

PR 17-OCT-1997; 97US-0062350P
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088302P.
PR 05-JUN-1998; 98US-0088312P.
PR 05-JUN-1998; 98US-0088317P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.

PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 18-JUN-1998; 98US-0089653P.
PR 19-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090423P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 02-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095986P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 15-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.

PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 64.3%; Score 18; DB 6; Length 4440;
Best Local Similarity 16.7%; Pred. No. 0.00011;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

Db 3858 CATTTTAATAC 3869

Search completed: January 4, 2006, 15:56:18
Job time : 83.1217 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 ; Search time 12.4174 Seconds
(without alignments)
92.983 Million cell updates/sec

Title: US-09-932-322-12
Perfect score: 28
Sequence: 1 CXXXXXXXXXC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 1000 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	64.3	307	1	GSFPR3
2	17	60.7	245	2	T24565
3	16	57.1	152	2	D72630
4	16	57.1	244	2	T31838
5	16	57.1	248	2	T31841
6	16	57.1	257	2	T21029
7	16	57.1	539	2	T52565
8	16	57.1	667	2	A48579
9	16	57.1	770	2	T51024
10	16	57.1	1041	2	T31097
11	15	53.6	154	2	AC0416
12	15	53.6	226	2	B41378
13	15	53.6	248	2	T21786
14	15	53.6	268	2	A84746
15	15	53.6	273	2	S69193
16	15	53.6	283	2	AH2744
17	15	53.6	287	2	S57770
18	15	53.6	289	2	S49812
19	15	53.6	318	2	G97525
20	15	53.6	357	2	T21152
21	15	53.6	363	2	S18765
22	15	53.6	392	2	T27303
23	15	53.6	409	2	T08928
24	15	53.6	420	2	S74388
25	15	53.6	423	2	D96552
26	15	53.6	423	2	S31126
27	15	53.6	569	2	B87180
28	15	53.6	727	2	UC7818
29	15	53.6	779	1	S40382

30	15	53.6	1003	1	PYVZAM
31	15	53.6	1043	2	T19734
32	15	53.6	1186	2	T18210
33	15	53.6	1364	2	T10236
34	15	53.6	1451	2	S65571
35	15	53.6	2098	2	T25888
36	15	53.6	2176	2	T13806
37	15	53.6	2187	2	S60224
38	15	53.6	2533	2	T28675
39	15	53.6	2533	2	T28675
40	15	53.6	2543	2	T31687
41	15	53.6	2703	1	A24420
42	15	53.6	4767	2	T31345
43	15	53.6	7829	2	T15789
44	15	53.6	7962	2	T18346
45	15	50.0	21	2	D22565
46	14	50.0	88	2	G82754
47	14	50.0	92	2	T54781
48	14	50.0	97	2	T28947
49	14	50.0	100	2	C72775
50	14	50.0	101	2	T14780
51	14	50.0	103	2	S73440
52	14	50.0	107	2	S46078
53	14	50.0	107	2	S57388
54	14	50.0	108	2	S67638
55	14	50.0	129	2	S48641
56	14	50.0	130	2	S57896
57	14	50.0	131	2	A43980
58	14	50.0	133	2	S44539
59	14	50.0	135	2	T15610
60	14	50.0	143	2	A54793
61	14	50.0	154	2	E72513
62	14	50.0	156	2	T31839
63	14	50.0	156	2	AH0865
64	14	50.0	156	2	C65065
65	14	50.0	158	2	S35201
66	14	50.0	159	2	E82692
67	14	50.0	164	2	T28741
68	14	50.0	165	2	T28742
69	14	50.0	166	2	T33970
70	14	50.0	177	2	S73288
71	14	50.0	177	2	S30932
72	14	50.0	177	2	F72706
73	14	50.0	195	2	S39967
74	14	50.0	197	2	T28739
75	14	50.0	199	2	S50363
76	14	50.0	204	2	T41853
77	14	50.0	205	2	T10384
78	14	50.0	209	2	JC7239
79	14	50.0	214	2	C70812
80	14	50.0	225	2	T22417
81	14	50.0	232	2	A60095
82	14	50.0	249	2	T31837
83	14	50.0	250	2	T31836
84	14	50.0	250	2	T31835
85	14	50.0	259	2	H71353
86	14	50.0	264	2	C70603
87	14	50.0	264	2	S42793
88	14	50.0	268	2	T05971
89	14	50.0	270	2	T27786
90	14	50.0	274	2	D82784
91	14	50.0	275	2	T05969
92	14	50.0	276	2	AE0860
93	14	50.0	286	1	C42053
94	14	50.0	288	2	T21158
95	14	50.0	293	2	T31840
96	14	50.0	298	2	S15524
97	14	50.0	300	2	T05729
98	14	50.0	300	2	S16344
99	14	50.0	312	2	T31834
100	14	50.0	315	2	T34526
101	14	50.0	328	2	T16065
102	14	50.0	329	2	T15823

spheroidin precurs
hypothetical prote
delta endotoxin -
xanthine dehydrog
pattern formation
hypothetical prote
lucan gene protei
polyketide synthas
alpha-SID immobili
surface antigen - P
notch protein - fr
hypothetical prote
hypothetical prote
elastic titlin - hu
R-phycocerythrin be
hypothetical prote
fibroblast growth
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
probable membrane
hypothetical prote
probable membrane
lysosome (BC 3.2.1
polyketide synthas
neurophysin 2 [val
probable membrane
hypothetical prote
pointiculin precurs
hypothetical prote
hypothetical prote
pepilin peptidase
pepilin peptidase
serine proteinase
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
nitrlase homolog
ACMMPV orf115 - BO
hypothetical prote
peroxiredoxin V -
probable lipQ prot
hypothetical prote
larval glue protei
hypothetical prote
hypothetical prote
hypothetical prote
conserved hypothet
hypothetical prote
PES-1 protein - Ca
probable receptor
phosphomethylpyrim
receptor-like kina
hypothetical prote
gap junction prote
hypothetical prote
hypothetical prote
variant surface gl
hypothetical prote
hypothetical prote
iron-sulfur cofact
hypothetical prote

249	13	46.4	66	2	A28644	kappa-neurotoxin -	322	13	46.4	201	2	T16181	hypothetical prote
250	13	46.4	66	2	JN0408	neurotoxin Os-1 -	323	13	46.4	202	2	C90742	hypothetical prote
251	13	46.4	71	2	T00008	copy number contro	324	13	46.4	202	2	F85592	hypothetical prote
252	13	46.4	79	2	T27605	hypothetical prote	325	13	46.4	204	2	D72864	AcOf-f-115 protein
253	13	46.4	81	2	B42465	probable transcrip	326	13	46.4	205	2	H71457	hypothetical prote
254	13	46.4	98	2	E84599	hypothetical prote	327	13	46.4	208	2	C97654	hypothetical prote
255	13	46.4	101	2	T42462	hypothetical prote	328	13	46.4	208	2	A12877	conserved hypotet
256	13	46.4	102	2	AE1357	hypothetical prote	329	13	46.4	209	2	AE2680	2-amino-4-hydroxy-
257	13	46.4	102	2	AE1727	hypothetical prote	330	13	46.4	210	2	S76973	hypothetical prote
258	13	46.4	102	2	T27988	hypothetical prote	331	13	46.4	213	2	F75585	urase accessory p
259	13	46.4	108	2	B95314	hypothetical prote	332	13	46.4	216	2	T50615	hypothetical prote
260	13	46.4	111	2	AC0733	probable exported	333	13	46.4	216	2	C90162	acetyltransferase,
261	13	46.4	111	2	T20452	hypothetical prote	334	13	46.4	217	2	A86600	CT734 hypothetical
262	13	46.4	111	2	T34565	hypothetical prote	335	13	46.4	217	2	A72024	conserved hypotet
263	13	46.4	112	2	T51142	nonspecific lipid-	336	13	46.4	217	2	A42644	neut protein - Esc
264	13	46.4	112	2	T48770	hypothetical prote	337	13	46.4	217	2	T48642	hypothetical prote
265	13	46.4	114	2	S57692	probable membrane	338	13	46.4	217	2	UC7875	L-rhamnose-binding
266	13	46.4	114	2	F69264	hypothetical prote	339	13	46.4	220	2	T25789	hypothetical prote
267	13	46.4	114	2	T01930	ethylene-response	340	13	46.4	222	2	T27939	hypothetical prote
268	13	46.4	117	2	T02169	homobox protein H	341	13	46.4	225	2	S45356	probable serine pr
269	13	46.4	119	2	B45937	early chorion prot	342	13	46.4	226	2	G95287	hypothetical prote
270	13	46.4	119	2	S24292	chorion protein -	343	13	46.4	226	2	T21380	hypothetical prote
271	13	46.4	119	2	S24294	chorion protein -	344	13	46.4	227	2	A33937	Ig light chain (13
272	13	46.4	119	2	S24291	chorion protein -	345	13	46.4	227	2	A55472	ornithine decarbox
273	13	46.4	121	2	S24293	mercury resistance	346	13	46.4	229	1	MMVYPM	25k protein - puta
274	13	46.4	125	2	S30170	probable resistance	347	13	46.4	233	2	T21496	hypothetical prote
275	13	46.4	126	2	AB0071	nucleic acid-bind	348	13	46.4	235	2	T16916	hypothetical prote
276	13	46.4	127	2	S27915	hypothetical prote	349	13	46.4	236	2	E84307	glucose 1-dehydrog
277	13	46.4	127	2	H90091	hypothetical prote	350	13	46.4	237	2	T25152	hypothetical prote
278	13	46.4	129	2	T47958	hypothetical prote	351	13	46.4	238	2	S77699	inner cell wall ma
279	13	46.4	132	1	MFIV2J	matrix protein M2	352	13	46.4	241	2	T22098	hypothetical prote
280	13	46.4	141	2	T33983	hypothetical prote	353	13	46.4	242	2	T24782	hypothetical prote
281	13	46.4	144	2	A36324	growth arrest-spec	354	13	46.4	242	2	A96794	unknown protein F1
282	13	46.4	145	2	H87436	hypothetical prote	355	13	46.4	243	2	H65136	hypothetical prote
283	13	46.4	145	2	T33037	hypothetical prote	356	13	46.4	243	2	F86006	hypothetical prote
284	13	46.4	146	1	W9BPT3	gene 19.2 protein	357	13	46.4	243	2	G91160	hypothetical prote
285	13	46.4	148	2	S19462	probable membrane	358	13	46.4	246	2	T00628	hypothetical prote
286	13	46.4	150	2	S66926	hypothetical prote	359	13	46.4	246	2	T25532	hypothetical prote
287	13	46.4	155	2	C82890	conserved hypotet	360	13	46.4	247	2	T21406	hypothetical prote
288	13	46.4	156	2	C91089	preillin peptidase	361	13	46.4	247	2	S72859	hypothetical prote
289	13	46.4	156	2	D75258	hypothetical prote	362	13	46.4	248	2	S05007	calcium channel al
290	13	46.4	156	2	E85934	preillin peptidase	363	13	46.4	248	2	T19913	hypothetical prote
291	13	46.4	157	2	JN0057	hypothetical 17K p	364	13	46.4	249	2	T21785	hypothetical prote
292	13	46.4	160	2	JN0503	peripheral myelin	365	13	46.4	249	2	T27782	hypothetical prote
293	13	46.4	160	2	S21721	growth arrest-spec	366	13	46.4	250	2	T36131	hypothetical prote
294	13	46.4	160	2	A41144	growth arrest-reia	367	13	46.4	250	2	T25725	hypothetical prote
295	13	46.4	162	2	T11978	phycoeryanin alpha	368	13	46.4	252	2	T25527	hypothetical prote
296	13	46.4	163	2	D83883	hypothetical prote	369	13	46.4	254	2	E83619	probable transport
297	13	46.4	163	2	B44813	probable molybdenu	370	13	46.4	255	2	T03408	MADS box protein -
298	13	46.4	166	1	NVBO2	vasopressin / neur	371	13	46.4	255	2	T34163	hypothetical prote
299	13	46.4	166	2	F81696	hypothetical prote	372	13	46.4	257	2	D71544	hypothetical prote
300	13	46.4	166	2	T32701	hypothetical prote	373	13	46.4	259	2	T27783	hypothetical prote
301	13	46.4	168	1	NVKT2	vasopressin / neur	374	13	46.4	261	2	S70006	finger protein zfo
302	13	46.4	168	2	B43755	vasopressin / neur	375	13	46.4	262	2	S01914	El membrane glycop
303	13	46.4	168	2	I38944	melanoma ubiquitin	376	13	46.4	262	2	G86938	conserved hypotet
304	13	46.4	171	2	H84923	hypothetical prote	377	13	46.4	265	2	T30059	hypothetical prote
305	13	46.4	173	2	P96543	hypothetical prote	378	13	46.4	266	2	T10816	thyroxine deiodina
306	13	46.4	173	2	A71144	hypothetical prote	379	13	46.4	268	2	T43486	hypothetical prote
307	13	46.4	176	2	T30961	hypothetical prote	380	13	46.4	269	2	T21407	hypothetical prote
308	13	46.4	177	2	A95283	lignin peroxidase	381	13	46.4	271	2	E86352	protein T26F17.16
309	13	46.4	180	2	PS0011	hypothetical prote	382	13	46.4	272	2	T19418	hypothetical prote
310	13	46.4	180	2	E71337	pollen-preferentia	383	13	46.4	274	2	T39087	T-cell surface pro
311	13	46.4	180	2	S29611	hypothetical prote	384	13	46.4	275	2	A34866	hypothetical prote
312	13	46.4	183	2	T38923	microfibril-associ	385	13	46.4	275	2	S08464	T-cell alloantigen
313	13	46.4	183	2	A54151	microfibril-associ	386	13	46.4	275	2	T05822	hypothetical prote
314	13	46.4	183	2	A49313	microfibril-associ	387	13	46.4	276	2	T43541	purine nucleotide
315	13	46.4	185	2	S57419	PAR-1a protein - c	388	13	46.4	276	2	AD1987	hypothetical prote
316	13	46.4	185	2	S62698	phococasinilate-re	389	13	46.4	277	2	H97266	mind family ATPase
317	13	46.4	185	2	S62699	phococasinilate-re	390	13	46.4	278	2	T16494	hypothetical prote
318	13	46.4	198	2	UB0316	proctozoa RNase, Dd	391	13	46.4	280	2	A42424	chitinase (EC 3.2.
319	13	46.4	199	2	T48099	hypothetical prote	392	13	46.4	287	1	A45168	probable 3-oxoacyl
320	13	46.4	199	2	T32267	hypothetical prote	393	13	46.4	287	2	T04236	xyloglucan endo-1,
321	13	46.4	201	2	S37847	hypothetical prote	394	13	46.4	289	2	AE2006	hypothetical prote

395	13	46.4	291	2	T28846	hypothetical prote
396	13	46.4	294	1	G65106	lipoprotein precu
397	13	46.4	294	2	ACO901	conserved hypothet
398	13	46.4	294	2	D91134	lipoprotein precu
399	13	46.4	294	2	G85979	lipoprotein precu
400	13	46.4	294	2	T23916	hypothetical prote
401	13	46.4	294	2	T23682	hypothetical prote
402	13	46.4	296	2	UT0548	Killer toxin KHR p
403	13	46.4	297	2	S06267	surface antigen H
404	13	46.4	300	2	T48147	hypothetical prote
405	13	46.4	301	2	T05493	pathogenesis-relat
406	13	46.4	301	2	T40593	cytoplasmic dynein
407	13	46.4	303	2	H71277	probable DNA aden
408	13	46.4	303	2	T46405	hypothetical prote
409	13	46.4	303	2	E88990	protein C36C5.13 f
410	13	46.4	309	2	S37934	hypothetical prote
411	13	46.4	311	2	H64483	5-methyltetrahydro
412	13	46.4	312	2	B75405	endopeptidase-rela
413	13	46.4	313	2	ACO603	asparaginase (EC 3
414	13	46.4	314	1	AS7126	dual specificity p
415	13	46.4	314	1	B57126	regenerating liver
416	13	46.4	317	2	T40050	probable elicitor
417	13	46.4	317	2	D86292	protein F7H2.14 f1
418	13	46.4	321	2	D64820	probable asparagin
420	13	46.4	323	2	S19650	cysteine proteinas
421	13	46.4	323	2	A39877	autocrine motility
422	13	46.4	324	1	D42951	alkanal monooxygen
423	13	46.4	324	2	S47432	cathepsin L (EC 3.
424	13	46.4	324	2	C86582	hypothetical prote
425	13	46.4	324	2	D72041	conserved hypothet
426	13	46.4	325	2	T27688	hypothetical prote
427	13	46.4	326	4	S61652	hypothetical prote
428	13	46.4	327	2	T41463	conserved hypothet
429	13	46.4	329	2	S74227	cathepsin K (EC 3.
430	13	46.4	330	2	T25323	hypothetical prote
431	13	46.4	335	2	JB0115	zinc-finger protei
432	13	46.4	339	2	P47301	Viril1 homolog - B
433	13	46.4	341	1	T22930	phosphoprotein pho
434	13	46.4	341	1	C70611	hypothetical prote
435	13	46.4	344	1	JO1643	hemagglutinin - in
436	13	46.4	346	1	MUHGCD	N4-Ibeta-N-acetylgl
437	13	46.4	347	2	S12955	calcium channel pr
438	13	46.4	348	2	T14141	NADH2 dehydrogenas
439	13	46.4	349	2	B81855	adenine glycosylas
440	13	46.4	349	2	D81088	A/G-specific adenl
441	13	46.4	353	2	T27800	hypothetical prote
442	13	46.4	354	2	S70595	NADH2 dehydrogenas
443	13	46.4	355	1	ADZW	fructose-bisphosph
444	13	46.4	355	2	TS1414	CONSTANS-like 1 -
445	13	46.4	358	2	T47550	fructose biphospho
446	13	46.4	358	2	A84781	probable fructose
447	13	46.4	361	2	S59531	manganese peroxida
448	13	46.4	362	2	T05824	probable pectin ac
449	13	46.4	363	2	JCS536	C-Fringe protein 1
450	13	46.4	364	1	TI8972	phosphoprotein pnc
451	13	46.4	364	1	C45574	40.7 kDa polypepti
452	13	46.4	364	2	H86384	40.1K Hypothetical
453	13	46.4	366	2	H89025	protein F13A2.6 f1
454	13	46.4	367	1	S24411	dual specificity p
455	13	46.4	367	1	S29090	dual specificity p
456	13	46.4	367	1	MOHU	alpha-2-HS-glycopr
457	13	46.4	367	2	AH0936	glycerol dehydroge
458	13	46.4	367	2	S52265	dual specificity p
459	13	46.4	367	2	D82763	UDP-N-acetylglucos
460	13	46.4	367	2	JC6087	helix-loop-helix t
461	13	46.4	368	2	H71600	refin PBB101ow - m
462	13	46.4	371	1	T07938	probable acetyl-Co
463	13	46.4	372	1	OPUGAP	lignin peroxidase
464	13	46.4	372	2	B32322	lignin peroxidase
465	13	46.4	372	2	JH0156	lignin peroxidase
466	13	46.4	372	2	A53050	breifeldin A estera
467	13	46.4	377	2	T04213	heat shock transcr
468	13	46.4	380	1	D65201	glycerol dehydroge
469	13	46.4	380	2	B91238	glycerol dehydroge
470	13	46.4	380	2	B66085	glycerol dehydroge
471	13	46.4	382	2	B86268	FL1B4.1 protein -
472	13	46.4	384	2	S25771	gaal protein - mou
473	13	46.4	389	2	AD0722	glutamate dehydrog
474	13	46.4	393	2	A56947	dual specificity p
475	13	46.4	393	2	A56115	dual specificity p
476	13	46.4	395	2	AB0814	probable alcohol
477	13	46.4	396	2	B49425	Desert hedgehog pr
478	13	46.4	400	2	B48613	surface protein Su
479	13	46.4	401	2	B87684	hypothetical prote
480	13	46.4	403	2	A13256	hypothetical prote
481	13	46.4	404	2	S58830	probable membrane
482	13	46.4	410	2	AB3546	aminobutyraldehyde
483	13	46.4	411	2	AB2760	cell division prot
484	13	46.4	412	2	T17653	hypothetical prote
485	13	46.4	412	2	C38104	Lfy floral meriste
486	13	46.4	415	2	C84829	hypothetical prote
487	13	46.4	415	2	T43352	nuclear receptor N
488	13	46.4	416	2	S65110	chitinase (EC 3.2.
489	13	46.4	420	2	B38104	Lfy floral meriste
490	13	46.4	422	2	S35197	hypothetical prote
491	13	46.4	422	2	A38104	Lfy floral meriste
492	13	46.4	425	2	B84631	probable serine ca
493	13	46.4	425	2	F70608	hypothetical prote
494	13	46.4	426	2	F85574	probable symport p
495	13	46.4	426	2	F90723	probable transport
496	13	46.4	430	2	JC4720	toxin co-regulated
497	13	46.4	433	1	S24353	proteaseome 266 sub
498	13	46.4	433	2	S53709	MS1 protein homol
499	13	46.4	433	2	S39349	MS1 protein homol
500	13	46.4	433	2	F86163	hypothetical prote
501	13	46.4	434	2	S50865	avermectin-sensiti
502	13	46.4	435	2	T20152	hypothetical prote
503	13	46.4	436	2	B81033	N-acetylglutamate
504	13	46.4	436	2	AB1977	probable amino-aci
505	13	46.4	437	2	S11497	branched-chain ami
506	13	46.4	437	2	H82295	branched chain ami
507	13	46.4	438	2	I38946	melanoma ubiqlitou
508	13	46.4	440	2	D69596	branched-chain ami
509	13	46.4	440	2	F97098	branched-chain ami
510	13	46.4	445	2	S37779	porin precursor, m
511	13	46.4	445	2	C68596	branched-chain ami
512	13	46.4	445	2	T18129	hypothetical prote
513	13	46.4	446	1	C70016	allantoinase homol
514	13	46.4	449	2	S50546	adenosylhomocyste
515	13	46.4	453	2	T09745	myb-related protei
516	13	46.4	454	2	D75446	oxidoreductase - D
517	13	46.4	455	2	T48973	cytochrome P450-1i
518	13	46.4	455	2	AC0308	isochoirismate gnt
519	13	46.4	457	2	I48206	triacylglycerol 1i
520	13	46.4	459	2	JC7931	cellobiohydrolase
521	13	46.4	461	2	S21223	triacylglycerol 1i
522	13	46.4	461	2	S50864	avermectin-sensiti
523	13	46.4	464	2	C46157	hexokinase (EC 2.7
524	13	46.4	465	2	A46157	hexokinase (EC 2.7
525	13	46.4	465	2	I49693	glucokinase (EC 2.
526	13	46.4	465	2	A31810	hexokinase (EC 2.7
527	13	46.4	465	2	S41084	triacylglycerol 1i
528	13	46.4	466	2	B46157	hexokinase (EC 2.7
529	13	46.4	467	2	E89605	protein P18G5.2 f1
530	13	46.4	471	2	S54460	hypothetical prote
531	13	46.4	472	1	A26730	ovotininbitor precu
532	13	46.4	474	2	B46746	glycine hydroxymet
533	13	46.4	474	2	AF0759	cobyrinic acid A,C
534	13	46.4	475	2	A33696	glycine hydroxymet
535	13	46.4	475	2	A71302	conserved hypothet
536	13	46.4	479	2	T49871	peroxisomal Ca-dep
537	13	46.4	482	2	T48370	S-receptor kinase
538	13	46.4	487	1	S71770	calcium-dependent
539	13	46.4	487	2	A43758	aromatic-L-amino-a
540	13	46.4	488	2	D70876	probable polyketid

541	13	46.4	488	2	T30602	hypothetical prote	614	13	46.4	650	2	S44806	F10B9.6 protein -
542	13	46.4	489	2	H69059	dihydroxy-acid den	615	13	46.4	654	2	D87791	protein B0207.12 l
543	13	46.4	494	2	C70940	probable cobo prot	616	13	46.4	656	2	D84359	thermosome subunit
544	13	46.4	495	2	T20754	hypothetical prote	617	13	46.4	656	2	E29826	hypothetical 72.4K
545	13	46.4	496	2	T33496	hypothetical prote	618	13	46.4	657	2	T52460	hypothetical prote
546	13	46.4	498	2	S12061	hexokinase (EC 2.7	619	13	46.4	659	2	A57532	Na+-dependent nucl
547	13	46.4	498	2	S27849	variant surface gl	620	13	46.4	663	2	T30621	hypothetical prote
548	13	46.4	504	2	T01603	hypothetical prote	621	13	46.4	664	2	S60062	hevin precursor -
549	13	46.4	504	2	E86343	T2211.12 protein	622	13	46.4	669	2	A46511	envelope protein -
550	13	46.4	504	2	T16526	hypothetical prote	623	13	46.4	676	2	G84663	hypothetical prote
551	13	46.4	506	2	S37583	RING finger protei	624	13	46.4	677	2	C42125	trochozoite cystei
552	13	46.4	506	2	T29368	hypothetical prote	625	13	46.4	684	2	G84730	Mutator-like trans
553	13	46.4	509	2	A95985	probable sugar kin	626	13	46.4	690	2	T11749	transferrin - Acta
554	13	46.4	511	2	A95985	cellulogamin I sylv	627	13	46.4	691	2	T31552	hypothetical prote
555	13	46.4	513	1	TVHURF	ret finger protein	628	13	46.4	700	2	S38426	chaperonin 60 proc
556	13	46.4	517	2	T19962	hypothetical prote	629	13	46.4	713	2	A35502	major surface-labe
557	13	46.4	520	2	A96922	beta-glucosidase h	630	13	46.4	730	2	A75486	hypothetical prote
558	13	46.4	530	2	UC7979	cellobiohydrolase	631	13	46.4	733	1	A46373	probable serine/ch
559	13	46.4	536	2	T24000	hypothetical prote	632	13	46.4	738	2	D86345	hypothetical prote
560	13	46.4	538	2	S52472	cell fusion protei	633	13	46.4	746	2	S52770	subtilisin-like pr
561	13	46.4	539	2	A23923	carboxylesterase (634	13	46.4	757	2	UC7519	subtilisin-like se
562	13	46.4	539	2	T46132	hypothetical prote	635	13	46.4	757	2	T16609	hypothetical prote
563	13	46.4	540	2	A31584	carboxylesterase (636	13	46.4	764	2	H83879	3-oxoacyl-(acyl-ca
564	13	46.4	541	2	T33583	hypothetical prote	637	13	46.4	768	2	B41029	integrin beta-8 ch
565	13	46.4	541	2	B82342	probable transcrip	638	13	46.4	773	1	JH0609	protein-tyrosine-p
566	13	46.4	543	2	T27000	hypothetical prote	639	13	46.4	775	2	T00962	hypothetical prote
567	13	46.4	546	2	T49931	hypothetical prote	640	13	46.4	781	2	S43534	integrin beta3 - c
568	13	46.4	549	2	JX0054	carboxylesterase (641	13	46.4	795	2	T34468	hypothetical prote
569	13	46.4	554	2	A39060	carboxylesterase (642	13	46.4	796	1	UC1285	protein-tyrosine-p
570	13	46.4	557	2	A47162	cholesterase B (B	643	13	46.4	799	1	A38308	integrin beta-5 ch
571	13	46.4	557	2	A48434	variant-specific s	644	13	46.4	802	1	A36065	protein-tyrosine-p
572	13	46.4	560	1	F69059	arginine-CrNA lig	645	13	46.4	807	2	C85025	hypothetical prote
573	13	46.4	561	2	S71597	carboxylesterase (646	13	46.4	814	2	T49207	receptor kinase-11
574	13	46.4	561	2	S62788	carboxylesterase (647	13	46.4	829	1	A47373	protein-tyrosine-p
575	13	46.4	561	2	C64459	dihydroxy-acid deh	648	13	46.4	832	2	UC9051	protein tyrosine p
576	13	46.4	562	2	A55281	carboxylesterase (649	13	46.4	840	2	S41218	HTR1 protein - yea
577	13	46.4	563	1	VCWMV7	env polypepten -	650	13	46.4	843	1	JQ2229	DNA-directed DNA p
578	13	46.4	564	2	S08199	cytochrome-c3 hydr	651	13	46.4	848	1	T02053	S-receptor kinase
579	13	46.4	565	2	S10367	carboxylesterase (652	13	46.4	852	2	S28415	guanine nucleotide
580	13	46.4	566	2	S19307	carboxylesterase (653	13	46.4	852	2	T28790	hypothetical prote
581	13	46.4	566	2	S69889	hemagglutinin prec	654	13	46.4	885	2	UC7898	3',5'-cyclic nucle
582	13	46.4	567	1	A41010	carboxylesterase (655	13	46.4	891	2	T21942	hypothetical prote
583	13	46.4	567	2	T08405	hexose transport p	656	13	46.4	892	2	A41697	nitrate assimilat
584	13	46.4	569	2	S50771	hexose transport p	657	13	46.4	895	2	T23191	hypothetical prote
585	13	46.4	571	1	S30253	GABA transport pro	658	13	46.4	907	2	T27317	hypothetical prote
586	13	46.4	573	2	T27671	hypothetical prote	659	13	46.4	917	1	S15885	hexokinase (EC 2.7
587	13	46.4	579	2	T30635	hypothetical prote	660	13	46.4	917	2	UC2025	hexokinase (EC 2.7
588	13	46.4	581	2	T24393	hypothetical prote	661	13	46.4	931	2	T25148	hypothetical prote
589	13	46.4	582	1	VCVDAR	env polypepten -	662	13	46.4	934	2	G86548	polymorphic outer
590	13	46.4	591	2	T48141	acroganin - guine	663	13	46.4	934	2	A72075	hypomorphlic prote
591	13	46.4	594	2	H48613	env polypepten pr	664	13	46.4	950	2	F86222	hypothetical prote
592	13	46.4	596	2	T30498	probable ribonucle	665	13	46.4	951	2	T00260	hypothetical prote
593	13	46.4	601	2	D89711	protein P40E10.4 (666	13	46.4	952	2	DB1593	polymorphic membra
594	13	46.4	601	2	T22025	hypothetical prote	667	13	46.4	962	2	UC5571	subtilisin-like pr
595	13	46.4	602	1	S25316	hydroxymethylgluta	668	13	46.4	966	2	B84053	subtilisin-like pr
596	13	46.4	602	2	T37254	acetylcholinestera	669	13	46.4	969	1	A39490	subtilisin-like pr
597	13	46.4	605	2	S48940	hypothetical prote	670	13	46.4	975	2	UC5570	subtilisin-like pr
598	13	46.4	609	2	UC7819	metalloprotease	671	13	46.4	983	2	T39902	translacion Elonga
599	13	46.4	616	2	T29234	hypothetical prote	672	13	46.4	993	2	S35633	DNA-binding protei
600	13	46.4	618	2	G82340	phosphoglucanate d	673	13	46.4	994	2	H96510	probable disease r
601	13	46.4	625	2	T10661	serine/chreonine-s	674	13	46.4	1011	2	T50344	poly(A)+ RNA trans
602	13	46.4	626	2	B97358	glucose-inhibited	675	13	46.4	1012	2	T52000	poly(A)+ RNA trans
603	13	46.4	632	2	S58152	hypothetical prote	676	13	46.4	1016	2	T00375	hypothetical prote
604	13	46.4	632	2	T21602	hypothetical prote	677	13	46.4	1046	2	A26838	prestalk protein p
605	13	46.4	633	2	T21898	hypothetical prote	678	13	46.4	1060	2	T30823	bunetanide sensiti
606	13	46.4	634	1	GERTX1	matrix glycoprotei	679	13	46.4	1073	2	B83051	cardamomylphosphate
607	13	46.4	635	2	S76371	hypothetical prote	680	13	46.4	1074	2	UC5928	semaphorin F precu
608	13	46.4	636	2	H87789	protein C34G6.1 (i	681	13	46.4	1082	2	T31112	ATPase 2 (EC 3.6.1
609	13	46.4	638	2	T26490	hypothetical prote	682	13	46.4	1087	2	S02035	period clock prote
610	13	46.4	640	1	HHKW7A	dnak-type molecula	683	13	46.4	1093	2	T50652	AP-3 complex beta3
611	13	46.4	640	2	T07923	acetyl-CoA carboxy	684	13	46.4	1094	2	T50651	AP3-complex beta-3
612	13	46.4	640	2	AB2251	glucose inhibited	685	13	46.4	1101	2	T16840	hypothetical prote
613	13	46.4	640	2	S62747	homeotic protein A	686	13	46.4	1105	2	C87622	Tomb-dependent rec

687	13	46.4	1112	2	H97050	DNA Polymerase III	760	13	46.4	2143	2	JH0427	voltage-dependent
688	13	46.4	1129	2	T42732	A-kinase anchoring	761	13	46.4	2143	2	G96595	hypothetical prote
689	13	46.4	1149	2	I38006	M130 antigen precu	762	13	46.4	2166	2	S11339	calcium channel pr
690	13	46.4	1151	2	I38004	M130 antigen precu	763	13	46.4	2168	2	T30171	nilein - mouse
691	13	46.4	1156	2	I38005	M130 antigen precu	764	13	46.4	2171	2	S05054	calcium channel al
692	13	46.4	1159	2	I38465	probable potassium	765	13	46.4	2180	2	T29764	hypothetical prote
693	13	46.4	1160	2	T13713	betal protein - fr	766	13	46.4	2188	2	A70984	probable polyketid
694	13	46.4	1162	2	T21557	hypothetical prote	767	13	46.4	2195	2	T34264	hypothetical prote
695	13	46.4	1166	2	T13958	synCAP-B1 protein	768	13	46.4	2201	2	S73014	polyketide synthas
696	13	46.4	1170	2	I45914	integrin alpha 2 s	769	13	46.4	2204	1	RANZNV	genome polyprotein
697	13	46.4	1176	2	T49482	hypothetical prote	770	13	46.4	2220	1	A45290	calcium channel pr
698	13	46.4	1180	2	A35854	integrin alpha-1 c	771	13	46.4	2233	2	T28669	surface protein 51
699	13	46.4	1181	2	A33998	integrin alpha-2 c	772	13	46.4	2240	2	T37057	probable mult-dom
700	13	46.4	1205	2	T41987	hypothetical prote	773	13	46.4	2257	2	D86483	protein f55.19 [i
701	13	46.4	1242	2	S14201	probable adenylate	774	13	46.4	2257	2	T09538	acetyl-CoA carboxy
702	13	46.4	1249	2	T14270	Ras-GRase activat	775	13	46.4	2261	2	T07084	acetyl-CoA carboxy
703	13	46.4	1265	2	S57968	Ran-binding protei	776	13	46.4	2304	2	T07920	probable acetyl-Co
704	13	46.4	1271	2	T24008	hypothetical prote	777	13	46.4	2359	2	E86483	probable acetyl-Co
705	13	46.4	1274	2	T10729	transferrin-like p	778	13	46.4	2518	2	A12140	polyketide synthas
706	13	46.4	1274	2	T42017	cysteine rich proc	779	13	46.4	2524	2	A35844	Korch protein - Af
707	13	46.4	1293	2	T14259	ras GRase-activat	780	13	46.4	2531	2	S18188	notch protein homo
708	13	46.4	1299	2	T43251	furin (EC 3.4.21.7	781	13	46.4	2531	2	A46019	notch-l protein -
709	13	46.4	1323	2	PN0568	connectin 3B - chl	782	13	46.4	2543	2	F69679	polyketide synthas
710	13	46.4	1323	2	E88257	protein let-23 [lm	783	13	46.4	2555	2	A40043	hypothetical prote
711	13	46.4	1324	2	S06187	RNA2 polypotein -	784	13	46.4	2610	2	T20968	variant-specific s
712	13	46.4	1331	2	S05011	calcium channel al	785	13	46.4	2706	2	T28155	hypothetical prote
713	13	46.4	1340	2	D96521	protein F21D18.16	786	13	46.4	2844	2	S28291	fibillin-2 precu
714	13	46.4	1359	2	T10235	xanthine dehydrog	787	13	46.4	2918	2	A54105	hypothetical prote
715	13	46.4	1369	2	T43433	alpha-glucan synth	788	13	46.4	2946	2	T15840	protein HMWPI - Ye
716	13	46.4	1374	2	S70712	protein-tyrosine k	789	13	46.4	3161	2	T30342	protein HMWPI - Ye
717	13	46.4	1391	2	T20406	hypothetical prote	790	13	46.4	3163	2	AB0233	Yersiniaabactin bio
718	13	46.4	1402	2	D70634	probable polyketid	791	13	46.4	3224	1	T17440	probable polyketid
719	13	46.4	1407	2	S59823	probable membrane	792	13	46.4	3224	1	S58884	Ran-binding protei
720	13	46.4	1416	2	E88550	protein ZC84.1 [lm	793	13	46.4	3461	2	S58870	reelin precursor
721	13	46.4	1446	2	S73013	polyketide synthas	794	13	46.4	3507	2	T34513	hypothetical prote
722	13	46.4	1461	2	T26327	hypothetical prote	795	13	46.4	3828	2	T13857	cithoxax protein
723	13	46.4	1484	2	T42632	breast cancer tumo	796	13	46.4	3971	2	T44806	mycosubtilin synth
724	13	46.4	1488	2	AG2136	polyketide synthas	797	13	46.4	4162	2	T42633	connectin/citin -
725	13	46.4	1488	2	C70984	probable ppsb prot	798	13	46.4	5376	2	T42215	zonadhesin - mouse
726	13	46.4	1489	2	S73015	polyketide synthas	799	13	46.4	6420	2	T30283	polyketide synthas
727	13	46.4	1538	2	E70874	probable ppsb prot	800	12	42.9	35	2	S55030	CAP5 protein - ant
728	13	46.4	1540	2	H87203	polyketide synthas	801	12	42.9	27	2	A59356	neurotoxin Bmk 41-
729	13	46.4	1565	2	AD2135	polyketide synthas	802	12	42.9	40	1	SMWF	metallothionein Mt
730	13	46.4	1567	2	T03730	antigen containing	803	12	42.9	40	2	B61194	metallothionein Mt
731	13	46.4	1570	2	AC2012	hypothetical prote	804	12	42.9	41	2	B44010	kit-ligand (altern
732	13	46.4	1587	2	AB2012	hypothetical prote	805	12	42.9	43	2	D84641	hypothetical prote
733	13	46.4	1597	1	BVPRSL	sol protein, large	806	12	42.9	45	1	XKPO2A	proteinasen inhibit
734	13	46.4	1597	2	T08428	gene small optic 1	807	12	42.9	46	2	S50015	leech-derived cryp
735	13	46.4	1722	2	E89753	protein P11C7.4 [l	808	12	42.9	50	2	F82539	hypothetical prote
736	13	46.4	1743	2	T26859	hypothetical prote	809	12	42.9	52	2	T10299	conotoxin-like pro
737	13	46.4	1774	2	T17421	hypothetical synthas	810	12	42.9	54	2	A31436	ovomucoid, thir d
738	13	46.4	1795	2	D97312	hypothetical prote	811	12	42.9	54	2	I61589	ovomucoid (PSTI-cy
739	13	46.4	1815	2	S73021	polyketide synthas	812	12	42.9	54	2	A61589	ovomucoid (PSTI-cy
740	13	46.4	1822	2	F87203	polyketide synthas	813	12	42.9	54	2	I61588	ovomucoid (PSTI-cy
741	13	46.4	1827	2	B70984	probable polyketid	814	12	42.9	54	2	F31439	ovomucoid, thir d
742	13	46.4	1844	1	RWPTM	genome polyprotein	815	12	42.9	54	2	I31446	ovomucoid, thir d
743	13	46.4	1844	1	S01956	hypothetical prote	816	12	42.9	54	2	C31447	ovomucoid, thir d
744	13	46.4	1871	2	AB7204	polyketide synthas	817	12	42.9	54	2	D61589	ovomucoid (PSTI-cy
745	13	46.4	1876	2	C70749	probable ppsa prot	818	12	42.9	54	2	G31442	ovomucoid, thir d
746	13	46.4	1959	1	AGRT	agrin - rat	819	12	42.9	54	2	A61492	ovomucoid (PSTI-cy
747	13	46.4	1964	2	T09059	notch4 - mouse	820	12	42.9	54	2	H31441	ovomucoid, thir d
748	13	46.4	1978	2	T07081	acetyl-CoA carboxy	821	12	42.9	54	2	H31442	ovomucoid, thir d
749	13	46.4	1984	2	T13171	probable vitellog	822	12	42.9	54	2	H31442	ovomucoid, thir d
750	13	46.4	1985	2	S19151	hypothetical prote	823	12	42.9	54	2	I31444	ovomucoid, thir d
751	13	46.4	2004	2	T30185	hypothetical prote	824	12	42.9	54	2	B31444	ovomucoid, thir d
752	13	46.4	2014	2	T21560	hypothetical prote	825	12	42.9	54	2	F31436	ovomucoid, thir d
753	13	46.4	2039	2	S60123	hypothetical prote	826	12	42.9	54	2	G61494	ovomucoid (PSTI-cy
754	13	46.4	2056	2	G88564	probable calcium c	827	12	42.9	54	2	H31445	ovomucoid, thir d
755	13	46.4	2066	2	G88564	protein R10B11.1 l	828	12	42.9	54	2	F61494	ovomucoid (PSTI-cy
756	13	46.4	2108	2	H70819	probable polyketid	829	12	42.9	54	2	F31447	ovomucoid, thir d
757	13	46.4	2124	2	T28658	polyketide synthas	830	12	42.9	54	2	F61492	ovomucoid (PSTI-cy
758	13	46.4	2126	2	E70522	probable polyketid	831	12	42.9	54	2	H61492	ovomucoid (PSTI-cy
759	13	46.4	2139	2	A44467	voltage-dependent	832	12	42.9	54	2	E31447	ovomucoid, thir d

833	12	42.9	54	2	B31443	ovomucoid, chird d	906	12	42.9	111	2	B85888	detox protein [imp
834	12	42.9	54	2	A31439	ovomucoid, chird d	907	12	42.9	111	2	UC7528	guanylate cyclase
835	12	42.9	54	2	G61492	ovomucoid (PSII-ty	908	12	42.9	112	1	XIHU	collipase precursor
836	12	42.9	54	2	H31436	ovomucoid, chird d	909	12	42.9	112	2	A46717	collipase precursor
837	12	42.9	54	2	B31436	ovomucoid, chird d	910	12	42.9	112	2	S64448	hypothetical prote
838	12	42.9	55	2	131442	ovomucoid, chird d	911	12	42.9	113	2	P95248	conserved hypotet
839	12	42.9	56	2	D31445	ovomucoid, chird d	912	12	42.9	113	2	C98113	conserved hypotet
840	12	42.9	56	2	A31446	ovomucoid, chird d	913	12	42.9	115	2	S66924	probable membrane
841	12	42.9	56	2	C61588	ovomucoid (PSII-ty	914	12	42.9	115	2	G95059	hypothetical prote
842	12	42.9	56	2	G98233	hypothetical prote	915	12	42.9	115	2	B96664	probable RING zinc
843	12	42.9	56	2	A25356	bdellin B-3 - medi	916	12	42.9	115	2	S45920	hypothetical prote
844	12	42.9	59	2	A47218	carboxyl ester lip	917	12	42.9	119	1	PSN33W	phospholipase A2 f
845	12	42.9	60	2	161886	zinc finger protei	918	12	42.9	121	1	QOEC13	hypothetical 13.5k
846	12	42.9	60	2	B86123	toxin III-8 - Tity	919	12	42.9	121	2	PH1661	Ig heavy chain V r
847	12	42.9	62	2	S62862	toxin III-8 - Tity	920	12	42.9	121	2	T02926	acyl carrier prote
848	12	42.9	62	2	S62866	toxin III-8 - Tity	921	12	42.9	121	2	A98089	hypothetical prote
849	12	42.9	62	2	B39510	metalothionein-2a	922	12	42.9	121	2	C85934	hypothetical prote
850	12	42.9	62	2	S54336	metalothionein-2c	923	12	42.9	121	2	UQ1058	DNA-binding protei
851	12	42.9	62	2	S54335	hypothetical prote	924	12	42.9	122	2	T18564	hypothetical prote
852	12	42.9	64	2	S66481	hypothetical prote	925	12	42.9	123	2	S30530	Ig heavy chain V r
853	12	42.9	66	2	E70083	hypothetical prote	926	12	42.9	124	2	T01251	hypothetical prote
854	12	42.9	67	2	JN0378	neurotoxin Os-3 -	927	12	42.9	125	1	NFB01	oxytocin / neuroph
855	12	42.9	67	2	S73050	probable acetyltra	928	12	42.9	125	1	NFBH1	oxytocin / neuroph
856	12	42.9	72	1	N2BPIJ	long neurotoxin 1	929	12	42.9	125	1	NFB01	oxytocin / neuroph
857	12	42.9	72	1	F82825	hypothetical prote	930	12	42.9	125	1	NFBT1	oxytocin / neuroph
858	12	42.9	73	1	N2OH2	long neurotoxin 2	931	12	42.9	125	1	NFSH1	oxytocin / neuroph
859	12	42.9	77	2	B82617	hypothetical prote	932	12	42.9	125	2	A43755	oxytocin / neuroph
860	12	42.9	78	2	E70891	hypothetical prote	933	12	42.9	128	2	S26786	Ig heavy chain V r
861	12	42.9	82	2	S70807	hypothetical prote	934	12	42.9	128	2	UQ2180	hypothetical 14.8k
862	12	42.9	84	1	UJ0152	acrosin/lypsin in	935	12	42.9	128	2	UQ1739	hypothetical 14.7k
863	12	42.9	85	2	D34770	ORF4 protein - sal	936	12	42.9	128	2	S27917	hypothetical prote
864	12	42.9	86	2	161885	T3B protein - rabb	937	12	42.9	129	2	S01661	lysosome (EC 3.2.1
865	12	42.9	89	2	D43692	proteinase inhibit	938	12	42.9	129	2	S07435	lysosome (EC 3.2.1
866	12	42.9	89	2	S71555	wheat aluminum ind	939	12	42.9	130	2	G71212	hypothetical prote
867	12	42.9	89	2	UQ2361	protein mcl - Mont	940	12	42.9	131	2	140656	hypothetical prote
868	12	42.9	92	2	T10783	probable chitinase	941	12	42.9	132	2	F64487	hypothetical prote
869	12	42.9	93	2	T06470	another-specific pr	942	12	42.9	133	2	T15955	hypothetical prote
870	12	42.9	94	2	T03285	collipase B precurs	943	12	42.9	133	2	T37311	hypothetical prote
871	12	42.9	96	1	XIHOB	keratin high-sulfu	944	12	42.9	133	2	JH0270	chondromodulin II
872	12	42.9	97	1	KRGTHM	keratin high-sulfu	945	12	42.9	133	2	JB0311	serine proteinase
873	12	42.9	97	1	KRSHM2	keratin high-sulfu	946	12	42.9	135	2	AF0865	conserved hypotet
874	12	42.9	97	2	T42335	hypothetical prote	947	12	42.9	135	2	S36197	hypothetical prote
875	12	42.9	97	2	UQ0041	hypothetical prote	948	12	42.9	135	2	G96691	hypothetical prote
876	12	42.9	97	2	D82789	hypothetical prote	949	12	42.9	136	2	T36241	lysosome (EC 3.2.1
877	12	42.9	97	2	C95984	hypothetical prote	950	12	42.9	137	1	JC4233	lysosome (EC 3.2.1
878	12	42.9	98	1	S25283	amoebapore A precu	951	12	42.9	137	2	T37213	hypothetical prote
879	12	42.9	98	1	KRSHM3	keratin high-sulfu	952	12	42.9	137	2	G84174	hypothetical prote
880	12	42.9	98	1	KRSHM4	keratin high-sulfu	953	12	42.9	138	2	AD0281	conserved hypotet
881	12	42.9	98	2	S26913	Ig heavy chain V r	954	12	42.9	138	2	S34725	regulation protein
882	12	42.9	98	2	147086	BIIB2 high-sulfur	955	12	42.9	139	2	T29955	hypothetical prote
883	12	42.9	98	2	T47086	hypothetical prote	956	12	42.9	140	2	JC5003	lysosome (EC 3.2.1
884	12	42.9	98	2	E96719	hypothetical prote	957	12	42.9	140	2	T49347	hypothetical prote
885	12	42.9	100	2	F95897	hypothetical prote	958	12	42.9	140	2	T05525	hypothetical prote
886	12	42.9	101	2	F75500	hypothetical prote	959	12	42.9	141	2	T33710	hypothetical prote
887	12	42.9	101	2	T26641	hypothetical prote	960	12	42.9	141	2	PC1294	trophozoite surfac
888	12	42.9	101	2	F72544	hypothetical prote	961	12	42.9	142	2	AC1934	protein B0294.1 [I
889	12	42.9	102	2	PH1249	Ig heavy chain V r	962	12	42.9	144	2	B89472	hypothetical prote
890	12	42.9	102	2	E89864	conserved hypotet	963	12	42.9	145	2	S74292	hemoglobin beta ch
891	12	42.9	102	2	S50530	hypothetical prote	964	12	42.9	145	2	S02027	protein [imported
892	12	42.9	102	2	A72604	hypothetical prote	965	12	42.9	145	2	G96581	conserved hypotet
893	12	42.9	104	2	150808	neuropeptide Y pre	966	12	42.9	145	2	T34539	hypothetical prote
894	12	42.9	104	2	T43968	orf 61.2 - phage T	967	12	42.9	146	2	B82568	conserved hypotet
895	12	42.9	104	2	F45681	conserved hypotet	968	12	42.9	146	2	E95363	thyrotropin beta c
896	12	42.9	106	2	F72043	CT657 hypothetical	969	12	42.9	147	2	S34148	Ig heavy chain var
897	12	42.9	106	2	C86580	hypothetical prote	970	12	42.9	148	2	147192	hypothetical prote
898	12	42.9	107	2	T51038	conserved plasmid	971	12	42.9	148	2	H71644	DNA-binding protei
899	12	42.9	107	2	A82642	hypothetical prote	972	12	42.9	148	2	B49727	hypothetical prote
900	12	42.9	108	2	E70931	conserved hypotet	973	12	42.9	148	2	T21334	helix-loop-helix p
901	12	42.9	108	2	AD0715	asparaginyllopepsi	974	12	42.9	149	2	JC5396	keratin high-sulfu
902	12	42.9	108	2	JN0368	prepilin-like proc	975	12	42.9	151	2	T36314	keratin high-sulfu
903	12	42.9	109	2	S54446	detox protein [imp	976	12	42.9	152	1	KRSHMC	cysteine proteinas
904	12	42.9	111	2	G91043	cca protein - Bsc	977	12	42.9	152	2	S57427	high-sulfur wool m
905	12	42.9	111	2	H65020		978	12	42.9	152	2	147111	

979 12 42.9 152 2 147109 high-sulfur wool m
980 12 42.9 152 2 147112 finger protein Hnf
981 12 42.9 152 2 C39384 hypothetical prote
982 12 42.9 153 2 C70958 hypothetical prote
983 12 42.9 153 2 A96751 hypothetical prote
984 12 42.9 154 2 S36584 B6 protein - human
985 12 42.9 154 2 A49727 DNA-binding protei
986 12 42.9 154 2 S47524 gene id1 protein-
987 12 42.9 155 1 T1ZM1 cryptin/factor XII
988 12 42.9 155 2 JC5395 helix-loop-helix p
989 12 42.9 155 2 D72761 hypothetical prote
990 12 42.9 155 2 JC7732 trypsin-plasmin in
991 12 42.9 157 2 S09805 hypothetical prote
992 12 42.9 158 2 B34332 vasococin / neurop
993 12 42.9 158 2 G86964 hypothetical prote
994 12 42.9 158 2 T26444 hypothetical prote
995 12 42.9 159 2 I84615 coagulation factor
996 12 42.9 159 2 JQ0136 hypothetical 17.8k
997 12 42.9 159 2 JQ1777 SalpR protein pre
998 12 42.9 159 2 T33596 SalpR protein pre
999 12 42.9 160 1 ERAD45 early E3 18.5k gly
1000 12 42.9 160 2 D48232 cysteine-rich exte

ALIGNMENTS

RESULT 1
GSFF3
salivary glue protein egs-3 - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster
C/Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 09-Jul-2004

C/Accession: A03329

R/Garlinkel, M.D.; Pruitt, R.E.; Meyerowitz, E.M.

J. Mol. Biol. 169, 765-789, 1983

A/Title: DNA sequences, gene regulation and modular protein evolution in the Drosophila

A/Reference number: A92904; MUID:83294545; PMID:6411930

A/Accession: A03329

A/Molecule type: DNA

A/Residues: 1-307 <GAR>

A/Cross-references: UNIPROT:P02840; UNIPARC:UPI00001245F4; GB:X01918; NID:98581; PIDN:CA

C/Comment: This protein is produced by third-instar larvae.

C/Genetics:

A/Genes: egs-3

A/Cross-references: FlyBase:FBgn0003373

A/Map position: 3L (68C)

A/Intons: 10/1

C/Superfamily: salivary glue protein

C/Keywords: salivary gland; tandem repeat

Query Match 64.3%; Score 18; DB 1; Length 307;
Best Local Similarity 16.7%; Pred. No. 1.3e-10;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 44 CTTTTTTTTTTC 55

RESULT 2
T24565
hypothetical protein T06C12.14 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T24565

R/Kelly, P.

submitted to the EMBL Data Library, October 1996

A/Reference number: Z19908

A/Accession: T24565

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-245 <WLL>

A/Cross-references: UNIPROT:O18039; UNIPARC:UPI000007C269; EMBL:Z81116; PIDN:CA803300.1;

A/Experimental source: clone T06C12

C/Genetics:

A/Genes: CESP:T06C12.14

A/Map position: 5

A/Intons: 33/3; 121/1

C/Superfamily: Caenorhabditis elegans hypothetical protein C4967.3

Query Match 60.7%; Score 17; DB 2; Length 245;
Best Local Similarity 16.7%; Pred. No. 5.8e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 193 CASATTTTSTC 204

RESULT 3
D72630
hypothetical protein APE1499 - Aeropyrum pernix (strain K1)

C/Species: Aeropyrum pernix

C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C/Accession: D72630

R/Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A/Reference number: A72450; MUID:99310339; PMID:10382966

A/Accession: D72630

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-152 <KAW>

A/Cross-references: UNIPROT:Q9YBV0; UNIPARC:UPI000005DFPF; DBJ:AF000061; NID:95104821;

A/Experimental source: strain K1

C/Genetics:

A/Genes: APE1499

Query Match 57.1%; Score 16; DB 2; Length 152;
Best Local Similarity 16.7%; Pred. No. 2.5e-07;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 137 CSTATTTTSTRC 148

RESULT 4
T31838

hypothetical protein T05B4.12 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C/Accession: T31838

R/Bradshaw, H.

submitted to the EMBL Data Library, July 1997

A/Description: The sequence of C. elegans cosmid T05B4.

A/Reference number: Z21092

A/Accession: T31838

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-244 <BRA>

A/Cross-references: UNIPROT:O16421; UNIPARC:UPI000007ACTE; EMBL:AF016445; PIDN:AA69052

A/Experimental source: strain Bristol N2; clone T05B4

C/Genetics:

A/Genes: CESP:T05B4.12

A/Map position: 5

A/Intons: 30/3; 116/1

C/Superfamily: Caenorhabditis elegans hypothetical protein C4967.3

Query Match 57.1%; Score 16; DB 2; Length 244;
Best Local Similarity 16.7%; Pred. No. 2.8e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

Db 203 CTSYADSSSTSC 214

RESULT 5

TJ1841

hypothetical protein T05B4.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C:Accession: TJ1841

R:Bradshaw, H.

submitted to the EMBL Data Library, July 1997

A:Description: The sequence of C. elegans cosmid T05B4.

A:Reference number: Z21092

A:Accession: TJ1841

A:Status: preliminary; translated from GB/EMBL/DDbU

A:Molecule type: DNA

A:Residues: 1-248

A:Cross-references: UNIPROT:O16424; UNIPARC:UPI000007F78C; EMBL:AF016445; PIDN:AAC69054.

A:Experimental source: strain Bristol N2; clone T05B4

C:Genetics:

A:Gene: CESP.T05B4.3

A:Map position: 5

A:Introns: 30/3; 118/1

C:Superfamily: Caenorhabditis elegans hypothetical protein C49G7.3

Query Match 57.1%; Score 16; DB 2; Length 248;

Best Local Similarity 16.7%; Pred. No. 2.8e-07;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

Db 207 CTSFADSSSSC 218

RESULT 6

TJ21029

hypothetical protein F16H6.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: TJ21029

R:Matthews, U.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19362

A:Accession: TJ21029

A:Status: preliminary; translated from GB/EMBL/DDbU

A:Molecule type: DNA

A:Residues: 1-257 <WIL>

A:Cross-references: UNIPROT:Q9XV90; UNIPARC:UPI0000075F8F; EMBL:Z81506; PIDN:CAB04129.1;

A:Experimental source: clone F16H6

C:Genetics:

A:Gene: CESP.F16H6.3

A:Map position: 5

A:Introns: 39/3; 123/1

C:Superfamily: Caenorhabditis elegans hypothetical protein C49G7.3

Query Match 57.1%; Score 16; DB 2; Length 257;

Best Local Similarity 16.7%; Pred. No. 2.8e-07;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

Db 195 CTNATTTTQASC 206

RESULT 7

T52565

glutathione synthase (EC 6.3.2.3) [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 09-Jul-2004

C:Accession: T52565

R:Skipsey, W.; Andrews, C.J.; Townsend, J.K.; Jepson, I.; Edwards, R.

submitted to the EMBL Data Library, July 1999

A:Description: Isolation of cDNA and genomic clones of glutathione synthetase containing

A:Reference number: Z26117

A:Accession: T52565

A:Status: preliminary; translated from GB/EMBL/DDbU

A:Molecule type: mRNA

A:Residues: 1-539 <SKI>

A:Cross-references: UNIPROT:P46416; UNIPARC:UPI000016DB47; EMBL:AJ243813; PIDN:CAB51027

A:Experimental source: cultivar Columbia; seedling; whole plant

C:Genetics:

A:Gene: gsh2

C:Superfamily: glutathione synthase

C:Keywords: ligase

Query Match 57.1%; Score 16; DB 2; Length 539;

Best Local Similarity 16.7%; Pred. No. 3.2e-07;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

Db 5 CSLSYSSSSSTC 16

RESULT 8

A48579

trichoxolite surface protein TSP11 - Giardia lamblia

C:Species: Giardia lamblia

C>Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: A48579

R:By, P.L.; Khanna, K.K.; Manning, P.A.; Mayrhofer, G.

MoI. Biochem. Parasitol. 58, 247-257, 1993

A:Title: A gene encoding a 69-kilodalton major surface protein of Giardia intestinalis

A:Reference number: A48579; PMID:8479449

A:Contents: Ad-1

A:Accession: A48579

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-667 <EY1>

A:Cross-references: UNIPROT:Q03185; UNIPARC:UPI0000137726; GB:M95814; NID:g159106; PIDN:

A:Note: sequence extracted from NCBI backbone (NCBIN:130056, NCBIPI:130058)

Query Match 57.1%; Score 16; DB 2; Length 667;

Best Local Similarity 16.7%; Pred. No. 3.4e-07;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

Db 449 CTTTARTVATC 460

RESULT 9

T51024

related to C2H2 zinc finger transcription factor D-Spl [imported] - Neurospora crassa

N/Alternate names: protein B7F21.50

C:Species: Neurospora crassa

C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004

C:Accession: T51024

R:Schulte, U.; Aism, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, July 2000

A:Reference number: Z25286

A:Accession: T51024

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-770 <SCH>

A:Cross-references: UNIPROT:Q9P319; UNIPARC:UPI000017B4E0; EMBL:AL389901; GSPDB:GN00116,

A:Experimental source: BAC clone B7F21; strain OR74A

C:Genetics:

A:Gene: NCSP:B7F21.50

A:Map position: 6

A:Introns: 117/1

Query Match 57.1%; Score 16; DB 2; Length 770;

Best Local Similarity 16.7%; Pred. No. 3.5e-07;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
Db 92 CSATTAAVGTAC 103

RESULT 10
T31097
chitin synthase (EC 2.4.1.16) CHS1 - fungus (*Fluobasidium floriforme*)
C/Species: *Fluobasidium neoformans*, *Cryptococcus neoformans*
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C/Accession: T31097
R/Species: C.A.
submitted to the EMBL Data Library, August 1997
A/Description: Chs1, a class IV chitin synthase of *Cryptococcus neoformans*.
A/Reference number: Z20980
A/Accession: T31097
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1041 <SPS>
A/Cross-references: UNIPROT:O13356; UNIPARC:UPI00001278CD; EMBL:AF021318; NID:g2444456;
A/Experimental source: strain H99
C/Genetics:
A/Gene: CHS1
A/Introns: 152/1; 556/1; 634/3; 922/2
C/Function:
A/Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucose
A/Pathway: chitin biosynthesis
A/Note: class IV chitin synthase
C/Superfamily: chitin synthase chs4
C/Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 57.1%; Score 16; DB 2; Length 1041;
Best Local Similarity 16.7%; Pred. No. 3.7e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
Db 212 CNSTSSSSGTC 223

RESULT 11
AC0416
probable prelinin peptidase dependent protein D precursor ppdd [imported] - *Yersinia pestis*
C/Species: *Yersinia pestis*
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C/Accession: AC0416
R/Parthill, J.; Wren, B.W.; Thomson, N.R.; Tilball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarrega, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.; Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
A/Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A/Reference number: AB0001; MUID:21470413; PMID:11586360
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-154 <RUR>
A/Cross-references: UNIPROT:Q8ZBIS; UNIPARC:UPI000000DC98A; GB:AL590842; PTDN:CAC92655.1;
C/Genetics:
A/Gene: ppdd
C/Superfamily: gonococcal fimbrial protein

Query Match 53.6%; Score 15; DB 2; Length 154;
Best Local Similarity 16.7%; Pred. No. 1.2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
Db 129 CTATDSAMDSC 140

RESULT 12
B41378
cytochrome c553i precursor - *Paracoccus denitrificans*

C/Species: *Paracoccus denitrificans*
C/Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 09-Jul-2004
C/Accession: B41378
R/Ras, J.; Reijnders, W.N.M.; Van Spanning, R.J.M.; Harms, N.; Olmann, L.F.; Stouthamer J. Bacteriol. 173, 6971-6979, 1991
A/Title: Isolation, sequencing, and mutagenesis of the gene encoding cytochrome c-553i o
A/Reference number: A41378; MUID:92041583; PMID:1657873
A/Accession: B41378
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-226 <RAS>
A/Cross-references: UNIPROT:P29967; UNIPARC:UPI0000126C5E; GB:M75583; NID:g150576; PTDN:
C/Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein
F/1-20/Domain: signal sequence #status predicted <SIG>
F/1-226/Product: cytochrome c553i #status predicted <MAT>
F/115-193/Domain: cytochrome c6 homology <CYC>
F/125,128/Binding site: heme (Cys) (covalent) #status predicted
F/129,173/Binding site: heme iron (His, Met) (axial ligands) #status predicted

Query Match 53.6%; Score 15; DB 2; Length 226;
Best Local Similarity 16.7%; Pred. No. 1.3e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
Db 16 CAASATAGTALC 27

RESULT 13
T21786
hypothetical protein F35E8.9 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T21786
R/Gardner, A.
submitted to the EMBL Data Library, November 1996
A/Reference number: Z19472
A/Accession: T21786
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-248 <WIL>
A/Cross-references: UNIPROT:O45450; UNIPARC:UPI0000077E4A; EMBL:281529; PTDN:CAE04292.1,
A/Experimental source: clone F35E8
C/Genetics:
A/Gene: CSR.F35E8.9
A/Map position: 5
A/Introns: 33/3; 121/1
C/Superfamily: *Caenorhabditis elegans* hypothetical protein C49G7.3

Query Match 53.6%; Score 15; DB 2; Length 248;
Best Local Similarity 16.7%; Pred. No. 1.3e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
Db 207 CSSYADSSTAC 218

RESULT 14
A84746
probable NAM (no apical meristem)-like protein [imported] - *Arabidopsis thaliana*
C/Species: *Arabidopsis thaliana* (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: A84746
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Talon, L. eus, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salberg, S.L.; Frazer, C.M.; Venter, J Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: A84746
A/Status: preliminary
A/Molecule type: DNA

A:Residues: 1-268 <STO>
A:Cross-references: UNIPROT:Q22798; UNIPARC:UPI00000A7D54; GB:AE002093; NID:92459430; P1
C:Genetics:
A:Gene: At2g33480
A:Map position: 2

Query Match 53.6%; Score 15; DB 2; Length 268;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 230 CCSSTSSSSVC 241

RESULT 15

S69193
Probable finger protein Pazf1 - garden pea
C:Species: Pisum sativum (garden pea)
C>Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S69193
R:Michael, A.J.; Hofer, J.M.I.; Ellis, T.H.N.
Plant Mol. Biol. 30, 1051-1058, 1996
A>Title: Isolation by PCR of a cDNA clone from pea petals with similarity to petunia and
A:Reference number: S69193; MUID:96270382; PMID:8639742
A:Accession: S69193
A:Molecule type: mRNA
A:Residues: 1-273 <MIC>
A:Cross-references: UNIPROT:Q41070; UNIPARC:UPI000009E1A7; EMBL:X87374; NID:9854399; P1D
C:Superfamily: Arabidopsis thaliana hypothetical protein F12E4.290
C:Keywords: DNA binding; zinc finger
F:115-135/Region: zinc finger CCH motif
F:202-222/Region: zinc finger CCH motif

Query Match 53.6%; Score 15; DB 2; Length 273;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 230 CTVANTTVTSSC 40

RESULT 16

AH2744
Conserved hypothetical protein Atu1368 [imported] - Agrobacterium tumefaciens (strain C5
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AH2744

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gilliet, W.; Grant, C.; Guenther, D.; Kutyavyn, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001

A:Author: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AH2744
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-283 <KUR>

A:Cross-references: UNIPROT:Q8UFM9; UNIPARC:UPI00001645DD; GB:AE008688; PIDN:AAL42374.1;
A:Experimental source: strain C58 (dupont)
C:Genetics:

A:Gene: Atu1368
A:Map position: circular chromosome

Query Match 53.6%; Score 15; DB 2; Length 283;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

DB 242 CSATHSDDMAAC 253

RESULT 17

S57770
xyloglucan endo-1,4-beta-D-glucanase (EC 3.2.1.-) precursor (clone CXET-B2) - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 06-Dec-1996 #sequence_revision 07-Feb-1997 #text_change 09-Jul-2004
C:Accession: S57770; S49813
R:Arrowsmith, D.A.; de Silva, J.

Plant Mol. Biol. 28, 391-403, 1995
A>Title: Characterisation of two tomato fruit-expressed cDNAs encoding xyloglucan endo-1
A:Reference number: S57769; MUID:95359399; PMID:7632911
A:Accession: S57770
A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-287 <ARR>

A:Cross-references: UNIPROT:Q43528; UNIPARC:UPI00000AAE06; EMBL:X82684; NID:9577067; P1
A>Note: the sequence from Fig. 2 is inconsistent with that from Fig. 5 in having 246-Gl
C:Superfamily: endoxyloglucan transferase

C:Keywords: glycosidase; hydrolase
F:1-18/Domain: signal sequence #status predicted <STG>
F:19-287/Product: xyloglucan endo-transglycosylase #status predicted <MAT>

Query Match 53.6%; Score 15; DB 2; Length 287;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 217 CWSAASATSTSSC 228

RESULT 18

S49812
xyloglucan endo-1,4-beta-D-glucanase (EC 3.2.1.-) precursor (clone CXET-B1) - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C:Accession: S57769; S49812
R:Arrowsmith, D.A.; de Silva, J.
Plant Mol. Biol. 28, 391-403, 1995
A>Title: Characterisation of two tomato fruit-expressed cDNAs encoding xyloglucan endo-1
A:Reference number: S57769; MUID:95359399; PMID:7632911
A:Accession: S57769
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-289 <AR2>

A:Cross-references: UNIPROT:Q43527; UNIPARC:UPI00000A4B68; EMBL:X82685; NID:9577065; P1L
C:Genetics:
A:Gene: CXET-B1
C:Superfamily: endoxyloglucan transferase

C:Keywords: glycosidase; hydrolase

Query Match 53.6%; Score 15; DB 2; Length 289;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 219 CWSAASATSTSSC 230

RESULT 19

G97525
CG10208 protein (AE003744) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: G97525

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001

A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: G97525
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-318 <KUR>
A:Cross-references: UNIPROT:Q8UFM9; UNIPARC:UPI00000D1B27; GB:AE007869; P1DN:AAK87160.1;
C:Genetics:
A:Gene: AGR_C_2528
A:Map position: circular chromosome

Query Match 53.6%; Score 15; DB 2; Length 318;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 277 CSATSSDMAAC 288

RESULT 20

T21152
hypothetical protein F20E11.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T21152
R:Ainscough, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19383
A:Accession: T21152
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-357 <MIT>
A:Cross-references: UNIPROT:Q9XV67; UNIPARC:UPI0000060FAD; EMBL:Z81508; P1DN:CAE04138.1;
A:Experimental source: clone F20E11
C:Genetics:
A:Gene: CESP:F20E11.7
A:Map position: 5
A:introns: 52/3; 88/2; 103/3; 154/3; 190/2; 205/3; 241/2; 270/3; 321/3

Query Match 53.6%; Score 15; DB 2; Length 357;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 7 CTATTGTMSAGC 18

RESULT 21

S18765
Sd protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S18765
R:Powers, P.A.; Ganetzky, B.
Genetics 129, 133-144, 1991
A:Title: On the components of segregation distortion in Drosophila melanogaster. V. Molecular
A:Reference number: S18765; MUID:92038937; PMID:1936954
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-363 <POW>
A:Cross-references: UNIPROT:P25722; UNIPARC:UPI000016BD5D; EMBL:X60218; NID:G11012; P1DN:
C:Genetics:
A:Gene: FlyBase:Sd
A:Cross-references: FlyBase:FBgn0024230

Query Match 53.6%; Score 15; DB 2; Length 363;
Best Local Similarity 16.7%; Pred. No. 1.4e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 331 CSRTSTFRSC 342

RESULT 22

T27303
hypothetical protein Y68A4A.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T27303
R:Steward, C.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z20340
A:Accession: T27303
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-392 <MIT>
A:Cross-references: UNIPROT:Q9XXP6; UNIPARC:UPI000007BED1; EMBL:AL021503; P1DN:CAA16424.
A:Experimental source: clone Y68A4A
C:Genetics:
A:Gene: CESP:Y68A4A.10
A:Map position: 5
A:introns: 45/2; 74/3; 139/3; 175/2; 190/3; 241/3; 305/3; 356/3

Query Match 53.6%; Score 15; DB 2; Length 392;
Best Local Similarity 16.7%; Pred. No. 1.5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 260 CTATTGTTSTGC 271

RESULT 23

T08928
sucrose cleavage protein homolog T15N24.70 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T08928
R:Bevan, M.; Zimmermann, W.; Gruenisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; May
submitted to the Protein Sequence Database, May 1999
A:Reference number: Z16518
A:Accession: T08928
A:Molecule type: DNA
A:Residues: 1-409 <BEV>
A:Cross-references: UNIPROT:Q9SUA2; UNIPARC:UPI000009E309; EMBL:AL078465; GSPDB:GN00062,
A:Experimental source: cultivar Columbia; BAC clone T15N24
C:Genetics:
A:Gene: ATSP:T15N24.70
A:Map position: 4
A:introns: 126/3; 155/2; 223/1; 299/3

Query Match 53.6%; Score 15; DB 2; Length 409;
Best Local Similarity 16.7%; Pred. No. 1.5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 394 CAAAYAVAYTC 405

RESULT 24

S74388
gamma-glutamyl phosphate reductase proA - Synecocystis sp. (strain PCC 6803)
N:Alternate names: hypothetical protein sl10461
C:Species: Synecocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S74388
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S74388
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-420 <RNX>
A:Cross-references: UNIPROT:Q55167; UNIPARC:UPI000013848; EMBL:U64001; GB:AB001339; NID
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: PROA
C:Superfamily: gamma-glutamyl phosphate reductase

Query Match 53.6%; Score 15; DB 2; Length 420;
Best Local Similarity 16.7%; Pred. No. 1.5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 153 CGASHSTATAC 164

RESULT 25
D96552
unknown protein, 74043-75895 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D96552
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hlizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luvo, J.S.; Maltl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96552
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-423 <STO>
A:Cross-references: UNIPROT:Q9C8U4; UNIPARC:UPI00000A997F; GB:AE05173; NID:G10092345; F
C:Genetics:
A:Gene: FSD21.9
A:Map position: 1
C:Superfamily: Arabidopsis thaliana hypothetical protein F24B22.230

Query Match 53.6%; Score 15; DB 2; Length 423;
Best Local Similarity 16.7%; Pred. No. 1.5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 306 CTTTSSLRASC 317

RESULT 26
S31126
hypothetical protein F59B2.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S31126
R:Julston, J.; Du, Z.; Thomas, K.; Wilson, R.; Hillier, L.; Staden, R.; Halloran, N.; G
awkins, T.; Alnecough, R.; Waterston, R.
submitted to the EMBL Data Library, November 1991
A:Description: The C. elegans sequencing project: A beginning.
A:Reference number: S31122
A:Accession: S31126
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-503 <SUL>
A:Cross-references: UNIPROT:P34482; UNIPARC:UPI000013B9D; EMBL:Z11505; NID:G6718; PID:G
C:Genetics:
A:introns: 26/3; 42/2; 121/3; 367/3; 443/1

C:Superfamily: Caenorhabditis elegans hypothetical protein F59B2.6

Query Match 53.6%; Score 15; DB 2; Length 503;
Best Local Similarity 16.7%; Pred. No. 1.5e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 4 CSASTSOLSTPFC 15

RESULT 27
B87180
pyruvate (or indolepyruvate) decarboxylase [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 05-Oct-2004
C:Accession: B87180
R:Coile, S.T.; Elgimeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H
R.; Davies, R.M.; Devlin, K.; Dutfoy, S.; Fellwell, T.; Fraser, A.; Hamlin, N.; Holroyd
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: B87180
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-569 <STO>
A:Cross-references: UNIPROT:Q9C8D6; UNIPARC:UPI00000C6E3D; GB:AL450380; NID:G13093715;
C:Genetics:
A:Gene: pdc
C:Superfamily: pyruvate decarboxylase/indolepyruvate decarboxylase; thiamin pyrophospha

Query Match 53.6%; Score 15; DB 2; Length 569;
Best Local Similarity 16.7%; Pred. No. 1.6e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 140 CSQANLTTATAC 151

RESULT 28
JC7818
metalloprotease I precursor - Alteromonas sp. (Strain O-7)
C:Species: Alteromonas sp. (Strain O-7)
C>Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 31-Dec-2004
C:Accession: JC7818
R:Myamoto, K.; Tsujibo, H.; Nukui, E.; Itoh, H.; Kaiznu, Y.; Inamori, Y.
Biochem. Biotechnol. Biochem. 66, 416-421, 2002
A:Title: Isolation and characterization of the genes encoding two metalloproteases (Mpr
A:Reference number: JC7818; PMID:11999419; MUID:21294062
A:Accession: JC7818
A:Molecule type: DNA
A:Residues: 1-727 <MTY>
A:Cross-references: UNIPROT:Q9VLQ5; UNIPARC:UPI000017CD49; DDBJ:AB063611
C:Comment: This enzyme, an extracellular alkaline metalloprotease, showing high similar
A:Gene: mpr1
A:Start codon: GTG
C:Superfamily: zinc metalloendopeptidase, neutral protease type (elastase)

Query Match 53.6%; Score 15; DB 2; Length 727;
Best Local Similarity 16.7%; Pred. No. 1.6e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 581 CNSTSTNSNC 592

RESULT 29

S40382
box A-binding factor - fruit fly (*Drosophila melanogaster*)
N:Alternate names: ABF; transcription factor GATrAb
C:Species: *Drosophila melanogaster*
C>Date: 13-Jan-1995 #sequence_revision 06-Sep-1996 #text_change 09-Jul-2004
C:Accession: S40382
R:Abel, T.; Michelson, A.M.; Maniatis, T.
Development 119, 623-633, 1993
A>Title: A *Drosophila* GATA family member that binds to Adh regulatory sequences is expressed in the developing embryo
A:Reference number: S40382; PMID:9424465; PMID:8187633
A:Accession: S40382
A:Molecule type: mRNA
A:Residues: 1-779 <ABF>
A:Cross-references: UNIPROT:P52172; UNIPARC:UPI0000135F5F; EMBL:X76217; NID:g441491; PID:G441491
C:Comment: This transcriptional activator is the earliest known marker of the developing embryo
C:Genetics:
A:Gene: FlyBase:FBgn0003507
A:Cross-references: FlyBase:FBgn0003507
C:Superfamily: box A-binding factor; GATA-type zinc finger homology
C:Keywords: DNA binding; nucleus; transcription factor; transcription regulation; zinc finger
F:316-369/Domain: GATA-type zinc finger homology <GZF>
F:319-343/Region: zinc finger GATA motif

Query Match 53.6%; Score 15; DB 1; Length 779;
Best Local Similarity 16.7%; Pred. No. 1.8e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 507 CSHSSNNRAC 518

RESULT 30

pyvZAM
spheroidin precursor - *Amsacta moorei* poxvirus
C:Species: *Amsacta moorei* poxvirus
C>Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: JQ1436; P00289; E41561
R:Banville, M.; Dunas, F.; Trifiro, S.; Arif, B.; Richardson, C.
J. Gen. Virol. 73, 559-566, 1992
A>Title: The predicted amino acid sequence of the spheroidin protein from *Amsacta moorei*
A:Reference number: JQ1436; PMID:92185464; PMID:1545219
A:Accession: JQ1436
A:Molecule type: DNA
A:Residues: 1-1003 <BAN>
A:Cross-references: UNIPROT:P29815; UNIPARC:UPI00000008A1; GB:M75889; NID:g209629; PIDN:A209629
A:Accession: P00289
A:Molecule type: protein
A:Residues: 2-7, 'X', 9-11, 'X', 13-86-90, 'X', 92-98, 'X', 100-102;532-540;728-750;786-802;810-812;814-816;818-820;822-824;826-828;830-832;834-836;838-840;842-844;846-848;850-852;854-856;858-860;862-864;866-868;870-872;874-876;878-880;882-884;886-888;890-892;894-896;898-900;902-904;906-908;910-912;914-916;918-920;922-924;926-928;930-932;934-936;938-940;942-944;946-948;950-952;954-956;958-960;962-964;966-968;970-972;974-976;978-980;982-984;986-988;990-992;994-996;998-1000;1002-1004;1006-1008;1010-1012;1014-1016;1018-1020;1022-1024;1026-1028;1030-1032;1034-1036;1038-1040;1042-1044;1046-1048;1050-1052;1054-1056;1058-1060;1062-1064;1066-1068;1070-1072;1074-1076;1078-1080;1082-1084;1086-1088;1090-1092;1094-1096;1098-1100;1102-1104;1106-1108;1110-1112;1114-1116;1118-1120;1122-1124;1126-1128;1130-1132;1134-1136;1138-1140;1142-1144;1146-1148;1150-1152;1154-1156;1158-1160;1162-1164;1166-1168;1170-1172;1174-1176;1178-1180;1182-1184;1186-1188;1190-1192;1194-1196;1198-1200;1202-1204;1206-1208;1210-1212;1214-1216;1218-1220;1222-1224;1226-1228;1230-1232;1234-1236;1238-1240;1242-1244;1246-1248;1250-1252;1254-1256;1258-1260;1262-1264;1266-1268;1270-1272;1274-1276;1278-1280;1282-1284;1286-1288;1290-1292;1294-1296;1298-1300;1302-1304;1306-1308;1310-1312;1314-1316;1318-1320;1322-1324;1326-1328;1330-1332;1334-1336;1338-1340;1342-1344;1346-1348;1350-1352;1354-1356;1358-1360;1362-1364;1366-1368;1370-1372;1374-1376;1378-1380;1382-1384;1386-1388;1390-1392;1394-1396;1398-1400;1402-1404;1406-1408;1410-1412;1414-1416;1418-1420;1422-1424;1426-1428;1430-1432;1434-1436;1438-1440;1442-1444;1446-1448;1450-1452;1454-1456;1458-1460;1462-1464;1466-1468;1470-1472;1474-1476;1478-1480;1482-1484;1486-1488;1490-1492;1494-1496;1498-1500;1502-1504;1506-1508;1510-1512;1514-1516;1518-1520;1522-1524;1526-1528;1530-1532;1534-1536;1538-1540;1542-1544;1546-1548;1550-1552;1554-1556;1558-1560;1562-1564;1566-1568;1570-1572;1574-1576;1578-1580;1582-1584;1586-1588;1590-1592;1594-1596;1598-1600;1602-1604;1606-1608;1610-1612;1614-1616;1618-1620;1622-1624;1626-1628;1630-1632;1634-1636;1638-1640;1642-1644;1646-1648;1650-1652;1654-1656;1658-1660;1662-1664;1666-1668;1670-1672;1674-1676;1678-1680;1682-1684;1686-1688;1690-1692;1694-1696;1698-1700;1702-1704;1706-1708;1710-1712;1714-1716;1718-1720;1722-1724;1726-1728;1730-1732;1734-1736;1738-1740;1742-1744;1746-1748;1750-1752;1754-1756;1758-1760;1762-1764;1766-1768;1770-1772;1774-1776;1778-1780;1782-1784;1786-1788;1790-1792;1794-1796;1798-1800;1802-1804;1806-1808;1810-1812;1814-1816;1818-1820;1822-1824;1826-1828;1830-1832;1834-1836;1838-1840;1842-1844;1846-1848;1850-1852;1854-1856;1858-1860;1862-1864;1866-1868;1870-1872;1874-1876;1878-1880;1882-1884;1886-1888;1890-1892;1894-1896;1898-1900;1902-1904;1906-1908;1910-1912;1914-1916;1918-1920;1922-1924;1926-1928;1930-1932;1934-1936;1938-1940;1942-1944;1946-1948;1950-1952;1954-1956;1958-1960;1962-1964;1966-1968;1970-1972;1974-1976;1978-1980;1982-1984;1986-1988;1990-1992;1994-1996;1998-2000;2002-2004;2006-2008;2010-2012;2014-2016;2018-2020;2022-2024;2026-2028;2030-2032;2034-2036;2038-2040;2042-2044;2046-2048;2050-2052;2054-2056;2058-2060;2062-2064;2066-2068;2070-2072;2074-2076;2078-2080;2082-2084;2086-2088;2090-2092;2094-2096;2098-2100;2102-2104;2106-2108;2110-2112;2114-2116;2118-2120;2122-2124;2126-2128;2130-2132;2134-2136;2138-2140;2142-2144;2146-2148;2150-2152;2154-2156;2158-2160;2162-2164;2166-2168;2170-2172;2174-2176;2178-2180;2182-2184;2186-2188;2190-2192;2194-2196;2198-2200;2202-2204;2206-2208;2210-2212;2214-2216;2218-2220;2222-2224;2226-2228;2230-2232;2234-2236;2238-2240;2242-2244;2246-2248;2250-2252;2254-2256;2258-2260;2262-2264;2266-2268;2270-2272;2274-2276;2278-2280;2282-2284;2286-2288;2290-2292;2294-2296;2298-2300;2302-2304;2306-2308;2310-2312;2314-2316;2318-2320;2322-2324;2326-2328;2330-2332;2334-2336;2338-2340;2342-2344;2346-2348;2350-2352;2354-2356;2358-2360;2362-2364;2366-2368;2370-2372;2374-2376;2378-2380;2382-2384;2386-2388;2390-2392;2394-2396;2398-2400;2402-2404;2406-2408;2410-2412;2414-2416;2418-2420;2422-2424;2426-2428;2430-2432;2434-2436;2438-2440;2442-2444;2446-2448;2450-2452;2454-2456;2458-2460;2462-2464;2466-2468;2470-2472;2474-2476;2478-2480;2482-2484;2486-2488;2490-2492;2494-2496;2498-2500;2502-2504;2506-2508;2510-2512;2514-2516;2518-2520;2522-2524;2526-2528;2530-2532;2534-2536;2538-2540;2542-2544;2546-2548;2550-2552;2554-2556;2558-2560;2562-2564;2566-2568;2570-2572;2574-2576;2578-2580;2582-2584;2586-2588;2590-2592;2594-2596;2598-2600;2602-2604;2606-2608;2610-2612;2614-2616;2618-2620;2622-2624;2626-2628;2630-2632;2634-2636;2638-2640;2642-2644;2646-2648;2650-2652;2654-2656;2658-2660;2662-2664;2666-2668;2670-2672;2674-2676;2678-2680;2682-2684;2686-2688;2690-2692;2694-2696;2698-2700;2702-2704;2706-2708;2710-2712;2714-2716;2718-2720;2722-2724;2726-2728;2730-2732;2734-2736;2738-2740;2742-2744;2746-2748;2750-2752;2754-2756;2758-2760;2762-2764;2766-2768;2770-2772;2774-2776;2778-2780;2782-2784;2786-2788;2790-2792;2794-2796;2798-2800;2802-2804;2806-2808;2810-2812;2814-2816;2818-2820;2822-2824;2826-2828;2830-2832;2834-2836;2838-2840;2842-2844;2846-2848;2850-2852;2854-2856;2858-2860;2862-2864;2866-2868;2870-2872;2874-2876;2878-2880;2882-2884;2886-2888;2890-2892;2894-2896;2898-2900;2902-2904;2906-2908;2910-2912;2914-2916;2918-2920;2922-2924;2926-2928;2930-2932;2934-2936;2938-2940;2942-2944;2946-2948;2950-2952;2954-2956;2958-2960;2962-2964;2966-2968;2970-2972;2974-2976;2978-2980;2982-2984;2986-2988;2990-2992;2994-2996;2998-3000;3002-3004;3006-3008;3010-3012;3014-3016;3018-3020;3022-3024;3026-3028;3030-3032;3034-3036;3038-3040;3042-3044;3046-3048;3050-3052;3054-3056;3058-3060;3062-3064;3066-3068;3070-3072;3074-3076;3078-3080;3082-3084;3086-3088;3090-3092;3094-3096;3098-3100;3102-3104;3106-3108;3110-3112;3114-3116;3118-3120;3122-3124;3126-3128;3130-3132;3134-3136;3138-3140;3142-3144;3146-3148;3150-3152;3154-3156;3158-3160;3162-3164;3166-3168;3170-3172;3174-3176;3178-3180;3182-3184;3186-3188;3190-3192;3194-3196;3198-3200;3202-3204;3206-3208;3210-3212;3214-3216;3218-3220;3222-3224;3226-3228;3230-3232;3234-3236;3238-3240;3242-3244;3246-3248;3250-3252;3254-3256;3258-3260;3262-3264;3266-3268;3270-3272;3274-3276;3278-3280;3282-3284;3286-3288;3290-3292;3294-3296;3298-3300;3302-3304;3306-3308;3310-3312;3314-3316;3318-3320;3322-3324;3326-3328;3330-3332;3334-3336;3338-3340;3342-3344;3346-3348;3350-3352;3354-3356;3358-3360;3362-3364;3366-3368;3370-3372;3374-3376;3378-3380;3382-3384;3386-3388;3390-3392;3394-3396;3398-3400;3402-3404;3406-3408;3410-3412;3414-3416;3418-3420;3422-3424;3426-3428;3430-3432;3434-3436;3438-3440;3442-3444;3446-3448;3450-3452;3454-3456;3458-3460;3462-3464;3466-3468;3470-3472;3474-3476;3478-3480;3482-3484;3486-3488;3490-3492;3494-3496;3498-3500;3502-3504;3506-3508;3510-3512;3514-3516;3518-3520;3522-3524;3526-3528;3530-3532;3534-3536;3538-3540;3542-3544;3546-3548;3550-3552;3554-3556;3558-3560;3562-3564;3566-3568;3570-3572;3574-3576;3578-3580;3582-3584;3586-3588;3590-3592;3594-3596;3598-3600;3602-3604;3606-3608;3610-3612;3614-3616;3618-3620;3622-3624;3626-3628;3630-3632;3634-3636;3638-3640;3642-3644;3646-3648;3650-3652;3654-3656;3658-3660;3662-3664;3666-3668;3670-3672;3674-3676;3678-3680;3682-3684;3686-3688;3690-3692;3694-3696;3698-3700;3702-3704;3706-3708;3710-3712;3714-3716;3718-3720;3722-3724;3726-3728;3730-3732;3734-3736;3738-3740;3742-3744;3746-3748;3750-3752;3754-3756;3758-3760;3762-3764;3766-3768;3770-3772;3774-3776;3778-3780;3782-3784;3786-3788;3790-3792;3794-3796;3798-3800;3802-3804;3806-3808;3810-3812;3814-3816;3818-3820;3822-3824;3826-3828;3830-3832;3834-3836;3838-3840;3842-3844;3846-3848;3850-3852;3854-3856;3858-3860;3862-3864;3866-3868;3870-3872;3874-3876;3878-3880;3882-3884;3886-3888;3890-3892;3894-3896;3898-3900;3902-3904;3906-3908;3910-3912;3914-3916;3918-3920;3922-3924;3926-3928;3930-3932;3934-3936;3938-3940;3942-3944;3946-3948;3950-3952;3954-3956;3958-3960;3962-3964;3966-3968;3970-3972;3974-3976;3978-3980;3982-3984;3986-3988;3990-3992;3994-3996;3998-4000;4002-4004;4006-4008;4010-4012;4014-4016;4018-4020;4022-4024;4026-4028;4030-4032;4034-4036;4038-4040;4042-4044;4046-4048;4050-4052;4054-4056;4058-4060;4062-4064;4066-4068;4070-4072;4074-4076;4078-4080;4082-4084;4086-4088;4090-4092;4094-4096;4098-4100;4102-4104;4106-4108;4110-4112;4114-4116;4118-4120;4122-4124;4126-4128;4130-4132;4134-4136;4138-4140;4142-4144;4146-4148;4150-4152;4154-4156;4158-4160;4162-4164;4166-4168;4170-4172;4174-4176;4178-4180;4182-4184;4186-4188;4190-4192;4194-4196;4198-4200;4202-4204;4206-4208;4210-4212;4214-4216;4218-4220;4222-4224;4226-4228;4230-4232;4234-4236;4238-4240;4242-4244;4246-4248;4250-4252;4254-4256;4258-4260;4262-4264;4266-4268;4270-4272;4274-4276;4278-4280;4282-4284;4286-4288;4290-4292;4294-4296;4298-4300;4302-4304;4306-4308;4310-4312;4314-4316;4318-4320;4322-4324;4326-4328;4330-4332;4334-4336;4338-4340;4342-4344;4346-4348;4350-4352;4354-4356;4358-4360;4362-4364;4366-4368;4370-4372;4374-4376;4378-4380;4382-4384;4386-4388;4390-4392;4394-4396;4398-4400;4402-4404;4406-4408;4410-4412;4414-4416;4418-4420;4422-4424;4426-4428;4430-4432;4434-4436;4438-4440;4442-4444;4446-4448;4450-4452;4454-4456;4458-4460;4462-4464;4466-4468;4470-4472;4474-4476;4478-4480;4482-4484;4486-4488;4490-4492;4494-4496;4498-4500;4502-4504;4506-4508;4510-4512;4514-4516;4518-4520;4522-4524;4526-4528;4530-4532;4534-4536;4538-4540;4542-4544;4546-4548;4550-4552;4554-4556;4558-4560;4562-4564;4566-4568;4570-4572;4574-4576;4578-4580;4582-4584;4586-4588;4590-4592;4594-4596;4598-4600;4602-4604;4606-4608;4610-4612;4614-4616;4618-4620;4622-4624;4626-4628;4630-4632;4634-4636;4638-4640;4642-4644;4646-4648;4650-4652;4654-4656;4658-4660;4662-4664;4666-4668;4670-4672;4674-4676;4678-4680;4682-4684;4686-4688;4690-4692;4694-4696;4698-4700;4702-4704;4706-4708;4710-4712;4714-4716;4718-4720;4722-4724;4726-4728;4730-4732;4734-4736;4738-4740;4742-4744;4746-4748;4750-4752;4754-4756;4758-4760;4762-4764;4766-4768;4770-4772;4774-4776;4778-4780;4782-4784;4786-4788;4790-4792;4794-4796;4798-4800;4802-4804;4806-4808;4810-4812;4814-4816;4818-4820;4822-4824;4826-4828;4830-4832;4834-4836;4838-4840;4842-4844;4846-4848;4850-4852;4854-4856;4858-4860;4862-4864;4866-4868;4870-4872;4874-4876;4878-4880;4882-4884;4886-4888;4890-4892;4894-4896;4898-4900;4902-4904;4906-4908;4910-4912;4914-4916;4918-4920;4922-4924;4926-4928;4930-4932;4934-4936;4938-4940;4942-4944;4946-4948;4950-4952;4954-4956;4958-4960;4962-4964;4966-4968;4970-4972;4974-4976;4978-4980;4982-4984;4986-4988;4990-4992;4994-4996;4998-5000;5002-5004;5006-5008;5010-5012;5014-5016;5018-5020;5022-5024;5026-5028;5030-5032;5034-5036;5038-5040;5042-5044;5046-5048;5050-5052;5054-5056;5058-5060;5062-5064;5066-5068;5070-5072;5074-5076;5078-5080;5082-5084;5086-5088;5090-5092;5094-5096;5098-5100;5102-5104;5106-5108;5110-5112;5114-5116;5118-5120;5122-5124;5126-5128;5130-5132;5134-5136;5138-5140;5142-5144;5146-5148;5150-5152;5154-5156;5158-5160;5162-5164;5166-5168;5170-5172;5174-5176;5178-5180;5182-5184;5186-5188;5190-5192;5194-5196;5198-5200;5202-5204;5206-5208;5210-5212;5214-5216;5218-5220;5222-5224;5226-5228;5230-5232;5234-5236;5238-5240;5242-5244;5246-5248;5250-5252;5254-5256;5258-5260;5262-5264;5266-5268;5270-5272;5274-5276;5278-52

A:Gene: ATSP:T1111.140
A:Map position: 4
A:Introns: 35/1; 66/2; 115/3; 613/3; 694/3; 799/3; 879/3; 974/3; 1048/3; 1102/3; 1147/3;
C:Superfamily: xanthine dehydrogenase; ferredoxin [2Fe-2S] homology
C:Keywords: 2Fe-2S; metalloprotein
F:43,48,51,86/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 53.6%; Score 15; DB 2; Length 1364;
Best Local Similarity 16.7%; Pred. No. 1.9e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 209 CGSKTTSEATC 220

RESULT 34
S65571
pattern formation protein GNOM - Arabidopsis thaliana
N:Alternate names: EMB30 protein
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: S65571; S65572
Mol. Gen. Genet. 250, 681-691, 1996
R:Busch, M.; Mayer, U.; Uerogens, G.
A>Title: Molecular analysis of the Arabidopsis pattern formation gene GNOM: gene structure
A:Reference number: S65571; MUID:96204508; PMID:8628228
A:Accession: S65571
A:Molecule type: DNA
A:Residues: 1-1451 <BUS>
A:Cross-references: UNIPROT:Q42510; UNIPARC:UPI0000129E82; EMBL:U56433; NID:g1209632; PI
A:Accession: S65572
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-110, 'T', 112-866, 'G', 868-1451 <BUS>
A:Cross-references: UNIPARC:UPI000016DA45; EMBL:U36432; NID:g1209630; PID:g1209631
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995
C:Genetics:
A:Gene: GNOM: EMB30
A:Introns: 246/3

Query Match 53.6%; Score 15; DB 2; Length 1451;
Best Local Similarity 16.7%; Pred. No. 1.9e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 22 CTDSNTTTLAC 33

RESULT 35
T25888
hypothetical protein T10H10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25888
R:Nelson, J.; Langston, Y.
Submitted to the EMBL Data Library, December 1996
A:Description: The sequence of C. elegans cosmid T10H10.
A:Reference number: Z20105
A:Accession: T25888
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2098 <NEL>
A:Cross-references: UNIPROT:P91443; UNIPARC:UPI00000756CE; EMBL:U80848; PID:AA837988.1;
A:Experimental source: strain Bristol N2; clone T10H10
C:Genetics:
A:Gene: CESP:T10H10.1
A:Map position: X
A:Introns: 6/3; 43/3; 92/3; 154/2; 226/1; 280/3; 417/2; 472/3; 812/3; 853/3; 887/3; 978/
3
F:65-720/Domain: myosin motor domain homology <MMO>

Query Match 53.6%; Score 15; DB 2; Length 2098;
Best Local Similarity 16.7%; Pred. No. 2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 1188 CADSATTAELC 1199

RESULT 36
T13806
toucan gene protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13806
R:Grammont, M.; Daatugue, B.; Couderc, J.L.
Development 124, 4917-4926, 1997
A>Title: The Drosophila toucan (toc) gene is required in germline cells for somatic cell
A:Reference number: Z17769; MUID:98090047; PMID:9362455
A:Accession: T13806
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2176 <GRA>
A:Cross-references: UNIPROT:Q46112; UNIPARC:UPI0000083F2C; EMBL:Y14157; NID:g2760521; P
A:Genetics:
A:Gene: toc
A:Cross-references: FlyBase:FBgn0015600

Query Match 53.6%; Score 15; DB 2; Length 2176;
Best Local Similarity 16.7%; Pred. No. 2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 2031 CTKASATTSATC 2042

RESULT 37
S60224
polyketide synthase 1 - Colletotrichum lagenarium
C:Species: Colletotrichum lagenarium
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S60224
R:Takano, Y.; Kubo, Y.; Shimizu, K.; Mise, K.; Okuno, T.; Furusawa, I.
Mol. Gen. Genet. 249, 162-167, 1995
A>Title: Structural analysis of PKS1, a polyketide synthase gene involved in melanin biosynthesis
A:Reference number: S60224; MUID:96086932; PMID:7500937
A:Accession: S60224
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-2187 <TKA>
A:Cross-references: UNIPROT:P79068; UNIPARC:UPI000006C308; GB:D83643; NID:g1208941; PIDN

Query Match 53.6%; Score 15; DB 2; Length 2187;
Best Local Similarity 16.7%; Pred. No. 2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 553 CSSAAALNVAC 564

RESULT 38
T28675
F:402-811/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
F:910-1198/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
F:166-1736/Domain: acyl carrier protein homology <ACP>
F:1788-1862/Domain: acyl carrier protein homology <ACPI>

alpha-51D immobilization antigen - Paramedium tetraurelia
C:Species: Paramedium tetraurelia
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T28675
R:Schwegmann, K.J.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z20506
A:Accession: T28675
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2533 <SCH>
A:Cross-references: UNIPROT:P90589; UNIPARC:UPI0000078A73; EMBL:X96400; PIDN:CAA65264.1
C:Gene: alpha-51D
A:Genetic code: SGC5
A:introns: 280/3; 538/2; 1248/2
C:Superfamily: G surface protein

Query Match 53.6%; Score 15; DB 2; Length 2533;
Best Local Similarity 16.7%; Pred. No. 2.1e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12
DB 360 CKGTSTTTEAAC 371

RESULT 39
T28674
alpha-51D-immobilization antigen - Paramedium tetraurelia
C:Species: Paramedium tetraurelia
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T28674
R:Schmidt, H.J.
submitted to the EMBL Data Library, March 1995
A:Reference number: Z20505
A:Accession: T28674
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2533 <SCH>
A:Cross-references: UNIPROT:Q27183; UNIPARC:UPI000007E3B2; EMBL:X85135; NID:G728634; PIDN:G728634
C:Genetics:
A:Genetic code: SGC5
A>Note: alpha-51D
C:Superfamily: G surface protein

Query Match 53.6%; Score 15; DB 2; Length 2533;
Best Local Similarity 16.7%; Pred. No. 2.1e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12
DB 360 CKGTSTTTEAAC 371

RESULT 40
T31687
surface antigen - Paramedium primaurelia
C:Species: Paramedium primaurelia
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: T31687
R:Bourgain-Guglielmetti, F.; Caron, F.
Journal of Bacteriol. Microbiol. 43, 303-314, 1996
A:Title: Molecular characterization of the D surface protein gene subfamily in Paramedium primaurelia
A:Reference number: Z21061; MUID:96313351; PMID:8768434
A:Accession: T31687
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2543 <BOU>
A:Cross-references: UNIPROT:P90649; UNIPARC:UPI000007A903; EMBL:X96616; NID:G1235576; PIDN:G1235576
C:Genetics:
A:Genetic code: SGC5
C:Superfamily: G surface protein

Query Match 53.6%; Score 15; DB 2; Length 2543;
Best Local Similarity 16.7%; Pred. No. 2.1e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12
DB 360 CKGTSTTTEAAC 371

RESULT 41
A24420
notch protein - fruit fly (Drosophila melanogaster)
N:Alternate names: neurogenic repetitive locus protein
C:Species: Drosophila melanogaster
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004
C:Accession: A24420; A24768; S09358; A05267
R:Kidd, S.; Kelley, M.R.; Young, M.W.
Mol. Cell. Biol. 6, 3094-3108, 1986
A:Reference number: A24420; MUID:87064624; PMID:3097517
A:Accession: A24420
A:Molecule type: DNA
A:Residues: 1-2703 <KID>
A:Cross-references: UNIPROT:P07207; UNIPARC:UPI000016BCC6; GB:K03508; NID:G157991; PIDN:G157991
R:Wharton, K.A.; Johansen, K.M.; Xu, T.; Artavanis-Tsakonas, S.
Cell 43, 567-581, 1985
A:Reference number: A24768; MUID:86079539; PMID:3935325
A:Accession: A24768
A:Molecule type: mRNA
A:Residues: 1-48, 'I', '50-118', 'R', '120-230', 'I', '232-256', 'N', '258-266', 'A', '268-872', 'R', '874-958',
A:Cross-references: UNIPARC:UPI0000173D1F
A>Note: the authors translated the codon ATC for residue 49 as Thr, ATT for residue 2044
R:autz, D.
Nucleic Acids Res. 17, 6463-6471, 1989
A:Title: Hypervariability of simple sequences as a general source for polymorphic DNA markers
A:Reference number: S09358; MUID:89385974; PMID:2780284
A:Accession: S09358
A:Molecule type: DNA
A:Residues: 2505-2551, 'QOOQ', '2552-2576', 'E', '2578-2604 <TN>
A:Cross-references: UNIPARC:UPI0000173D20
R:Wharton, K.A.; Yedvobnick, B.; Flannery, V.G.; Artavanis-Tsakonas, S.
Cell 40, 55-62, 1985
A:Title: opa: a novel family of transcribed repeats shared by the Notch locus and other genes
A:Reference number: A05267; MUID:85099329; PMID:2981631
A:Accession: A05267
A:Molecule type: DNA
A:Residues: 2504-2576, 'E', '2578-2611 <WHA>
A:Cross-references: UNIPARC:UPI0000173D21
C:Genetics:
A:Gene: notch; opa
A:Cross-references: FlyBase:FBgn0004647
A:Map position: 8.96-9.36
A:introns: 53/3; 84/3; 171/3; 240/3; 283/3; 233/3; 2436/3; 2588/3
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology
C:Keywords: differentiation; tandem repeat; transmembrane protein
F:27-43/Domain: transmembrane #status predicted <TM1>
F:297-328/Domain: EGF homology <EGX1>
F:530-561/Domain: EGF homology <EGF1>
F:568-599/Domain: EGF homology <EGF>
F:988-1019/Domain: EGF homology <EGX2>
F:1064-1095/Domain: EGF homology <EGF3>
F:1187-1218/Domain: EGF homology <EGX3>
F:1746-1762/Domain: transmembrane #status predicted <TM2>
F:1950-1982/Domain: ankyrin repeat homology <AN1>
F:1988-2015/Domain: ankyrin repeat homology <AN2>
F:1988-2004/Domain: transmembrane #status predicted <TM3>
F:2017-2049/Domain: ankyrin repeat homology <AN3>
F:2050-2082/Domain: ankyrin repeat homology <AN4>
F:2083-2115/Domain: ankyrin repeat homology <AN5>
F:2538-2568/Region: glutamine-rich
F:2538-2568/Domain: neurogenic repetitive element #status predicted <OPA>

Query Match 53.6%; Score 15; DB 1; Length 2703;

Best Local Similarity 16.7%; Pred. No. 2.1e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

DB 192 CTALAGSSSFTC 203

RESULT 42

TJ31345

hypothetical protein G01D9.5 - *Caenorhabditis briggsae*C:Species: *Caenorhabditis briggsae*

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C:Accession: TJ31345

R:Waterston, R.

submitted to the EMBL Data Library, April 1996

A:Description: The C. briggsae genome sequencing project.

A:Reference number: Z21010

A:Accession: TJ31345

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4767 <MAT>

A:Cross-references: UNIPROT:Q17301; UNIPARC:UPI000011013B; EMBL:U56248; NID:G1233789; PI

C:Genetics:

A:Note: G01D9.5

A:Introns: 269/3; 341/3; 853/1; 920/2; 4452/3; 4534/3; 4592/3; 4654/3; 4670/3; 4707/2; 4

C:Keywords: carrier protein; phosphopantetheine; phosphoprotein

F:J472-3541/Domain: acyl carrier protein homology <ACP1>

F:4039-4427/Domain: acetate-CoA ligase homology <ACL>

F:4447-4514/Domain: acyl carrier protein homology <ACP2>

F:2210,3505/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 53.6%; Score 15; DB 2; Length 4767;
Best Local Similarity 16.7%; Pred. No. 2.4e-05;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

DB 1423 CSSSFHALSSAC 1434

RESULT 43

T15789

hypothetical protein C41A3.1 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C:Accession: T15789

R:Bentley, D.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid C41A3.

A:Reference number: Z18404

A:Accession: T15789

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-7829 <BEN>

A:Cross-references: UNIPROT:Q18559; UNIPARC:UPI000011014C; EMBL:U41541; NID:G1109867; PI

C:Genetics:

A:Gene: CESP_C41A3.1

A:Introns: 29/2; 69/3; 141/2; 192/2; 271/3; 321/2; 1000/1; 1098/1; 1127/3; 1168/2; 1282/

/2; 7514/3; 7596/3; 7654/3; 7716/3; 7732/3; 7769/2; 7797/3

C:Keywords: carrier protein; phosphopantetheine; phosphoprotein

F:6547-6616/Domain: acyl carrier protein homology <ACP>

F:2832,5521,6580/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 53.6%; Score 15; DB 2; Length 7829;
Best Local Similarity 16.7%; Pred. No. 2.6e-05;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

DB 4475 CSSSFYALSSAC 4486

RESULT 44
I38346
elastic titin - human (fragment)

C:Species: *Homo sapiens* (man)

C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004

C:Accession: I38346

R:Labelt, S.; Kolmerer, B.

Science 270, 293-296, 1995

A>Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.

A:Reference number: A57430; MUID:96026330; PMID:7569978

A:Accession: I38346

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-7962 <RES>

A:Cross-references: UNIPROT:Q10465; UNIPARC:UPI000011010C; EMBL:X90569; NID:G1017426; P

C:Genetics:

A:Gene: GDB:TTN

A:Cross-references: GDB:127867; OMIM:188840

A:Map position: 2q31-2q31

Query Match 53.6%; Score 15; DB 2; Length 7962;
Best Local Similarity 16.7%; Pred. No. 2.6e-05;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

DB 381 CTASNEAGSSSC 392

RESULT 45

D22565

R-phycoerythrin beta-2 chain - red alga (*Gastrocyclonium coulteri*) (fragment)C:Species: *Gastrocyclonium coulteri*

C>Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 09-Jul-2004

C:Accession: D22565

R:Klotz, A.V.; Glazer, A.N.

J. Biol. Chem. 260, 4856-4863, 1985

A>Title: Characterization of the bilin attachment sites in R-phycoerythrin.

A:Reference number: A22565; MUID:85182601; PMID:3886644

A:Accession: D22565

A:Molecule type: protein

A:Residues: 1-21 <KLO>

A:Cross-references: UNIPROT:Q7M276; UNIPARC:UPI0000174E0D

C:Superfamily: phycocyanin

Query Match 50.0%; Score 14; DB 2; Length 21;
Best Local Similarity 16.7%; Pred. No. 0.0004;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

DB 8 CMLSSVASYC 19

RESULT 46

G82754

hypothetical protein XF0861 [imported] - *Xylella fastidiosa* (strain 9a5c)C:Species: *Xylella fastidiosa*

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C:Accession: G82754

C:Anonymous: The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A>Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: For a complete list of authors see reference number A59328 below

A:Accession: G82754

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-88 <SIM>

A:Cross-references: UNIPROT:Q9PF17; UNIPARC:UPI00000C254B; GB:AE003925; GB:AE003849; NII

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;

Birones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carraro, D.M.; Carter, I

as-Neto, E.; Docena, C.; El-Dorfy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Prohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laizy
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328

A:Contents: annotation

A:Genetics:

A:Gene: XF0861

Query Match 50.0%; Score 14; DB 2; Length 88;
Best Local Similarity 16.7%; Pred. No. 0.00052;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12
Db 23 CSSSVTLKSSRC 34

RESULT 47

154781
fibroblast growth factor receptor FGFR-1, secreted splice form - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C:Accession: I54781
R:Leung, H.Y.; Gullick, W.J.; Lemoine, N.R.
Int. J. Cancer 59, 667-675, 1994
A:Title: Expression and functional activity of fibroblast growth factors and their recep
A:Reference number: I54781; MUID:95048906; PMID:7960240
A:Accession: I54781
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-92 <RES>
A:Cross-references: UNIPARC:UPI0000116DB9; GB:S76658; NID:9913557; PIDN:AAD14230.1; PID:
A:Experimental source: pancreatic cancer cell line PT45
C:Genetics:

A:Gene: GDB:FGFR1; FLT2
A:Cross-references: GDB:119913; OMIM:136350
A:Map position: 8p11.2-8p11.1
A:Introns: 58/3
C:Keywords: growth factor receptor

Query Match 50.0%; Score 14; DB 2; Length 92;
Best Local Similarity 16.7%; Pred. No. 0.00053;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12
Db 30 CGAGSASSSLC 41

RESULT 48

T28947
hypothetical protein F07C4.11 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T28947
R:Miller, N.; Steillyes, L.
submitted to the EMBL Data Library, January 1997
A:Description: The sequence of C. elegans cosmid F07C4.
A:Reference number: Z20546
A:Accession: T28947
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-97 <MTL>
A:Cross-references: UNIPROT:P91216; UNIPARC:UPI000007C37D; EMBL:U80023; PIDN:AAC48019.1;
A:Experimental source: strain Bristol N2, clone F07C4
C:Genetics:

A:Gene: CESP:F07C4.11
A:Map position: 5

Query Match 50.0%; Score 14; DB 2; Length 97;
Best Local Similarity 16.7%; Pred. No. 0.00053;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12
Db 56 CTTVAADSSPSC 67

RESULT 49

C72775
hypothetical protein APE0189 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: C72775
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: C72775
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100 <KAW>
A:Cross-references: UNIPROT:Q9YFR0; UNIPARC:UPI000005DA3D; DBD:AP000058; NID:95103388;
A:Experimental source: strain K1
C:Genetics:

A:Gene: APE0189
C:Superfamily: Aeropyrum pernix hypothetical protein APE0189

Query Match 50.0%; Score 14; DB 2; Length 100;
Best Local Similarity 16.7%; Pred. No. 0.00054;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12
Db 65 CASVSRMRASSC 76

RESULT 50

T14780
hypothetical protein DKFZ566B1346.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14780
R:Ottmawaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A:Reference number: Z18184
A:Accession: T14780
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-101 <OTT>
A:Cross-references: UNIPARC:UPI000016AC6B; EMBL:AL110253
A:Experimental source: fetal kidney; clone DKFZ566B1346
C:Genetics:

A:Note: DKFZ566B1346.1

Query Match 50.0%; Score 14; DB 2; Length 101;
Best Local Similarity 16.7%; Pred. No. 0.00054;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12
Db 38 CSISVVTGTASC 49

Search completed: January 4, 2006, 16:10:30
Job time: 13.4174 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 ; Search time 78.3652 Seconds
(without alignments)
108.037 Million cell updates/sec

Title: US-09-932-322-12
Perfect score: 28
Sequence: 1 CXXXXXXXXXXC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 10%
Listing first 1000 summaries

Database : UniProt 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	64.3	307	1	SGS3 DROME
2	18	64.3	347	2	Q8K340 MOUSE
3	18	64.3	362	2	Q8C0D3 MOUSE
4	18	64.3	367	2	Q6AV38 ORYSA
5	18	64.3	584	2	Q8BL43 MOUSE
6	18	64.3	730	2	Q8LPI3 PEA
7	17	60.7	35	2	Q9LQ12 ARATH
8	17	60.7	137	2	Q24960 GIALLA
9	17	60.7	137	2	Q24987 GIALLA
10	17	60.7	210	2	Q7XHS1 ORYSA
11	17	60.7	234	2	Q9XY90 GIALLA
12	17	60.7	245	2	Q18039 CAENORHABDI
13	17	60.7	306	2	Q5MQW8 CRYNNE
14	17	60.7	378	2	Q6ASQ2 ORYSA
15	17	60.7	396	2	Q4IGT2 9BURYK
16	17	60.7	429	2	Q8CX11 MOUSE
17	17	60.7	540	2	Q93H33 STRAW
18	17	60.7	667	2	Q9XFK3 GIALLA
19	17	60.7	1031	2	Q5B3H3 EMBENI
20	17	60.7	5703	1	MUC5B HUMAN
21	16	57.1	93	2	Q6EUA6 ORYSA
22	16	57.1	100	2	Q8S0L9 ORYSA
23	16	57.1	100	2	Q62YJ1 BURMA
24	16	57.1	110	1	CBP1 AJUCA
25	16	57.1	144	2	Q5C209 SCUA
26	16	57.1	152	2	Q9YBV0 ABRPE
27	16	57.1	161	2	Q7SLM9 ORYSA
28	16	57.1	172	2	Q7PY44 ANOGA
29	16	57.1	176	2	Q6ZD90 ORYSA
30	16	57.1	187	2	Q55014 DICDI
31	16	57.1	208	2	Q82LKO STRAW

32	16	57.1	210	2	Q67V9 ORYSA
33	16	57.1	215	2	Q5K9J0 CRYNNE
34	16	57.1	214	2	Q5SK09 CRYNNE
35	16	57.1	215	2	Q5K919 CRYNNE
36	16	57.1	227	2	Q4OSJ9 GLECA
37	16	57.1	230	2	Q7ONB5 GIALLA
38	16	57.1	234	2	Q8LD53 ARATH
39	16	57.1	234	2	Q9F3J7 ARATH
40	16	57.1	244	2	Q16421 CAEL
41	16	57.1	247	2	Q966B6 CAEL
42	16	57.1	248	2	Q16424 CAEL
43	16	57.1	257	2	Q9XV90 CAEL
44	16	57.1	262	2	Q8CAZ5 MOUSE
45	16	57.1	269	2	Q7PRQ7 ANOGA
46	16	57.1	278	2	Q6ZUB7 HUMAN
47	16	57.1	285	2	Q8PFE6 XANAC
48	16	57.1	286	2	Q6ZQJ1 HUMAN
49	16	57.1	287	2	Q9F934 BIPAD
50	16	57.1	297	2	Q89DY5 BRAJA
51	16	57.1	305	2	Q5OLQ5 ORYSA
52	16	57.1	309	2	Q5ZBR6 ORYSA
53	16	57.1	320	2	Q8N4O3 HUMAN
54	16	57.1	325	2	Q91E43 9ALPH
55	16	57.1	329	2	Q7WBR6 BORBA
56	16	57.1	329	2	Q7WPO8 BORBR
57	16	57.1	353	2	Q8XQX5 RALSO
58	16	57.1	366	2	Q55057 CRYNNE
59	16	57.1	381	2	Q5ZPV8 ORYSA
60	16	57.1	381	2	Q7Y021 ORYSA
61	16	57.1	426	2	Q941V9 ORYSA
62	16	57.1	437	2	Q7XUJ0 ORYSA
63	16	57.1	448	2	Q8BP04 MOUSE
64	16	57.1	465	2	Q7S1H3 NEUCR
65	16	57.1	471	2	Q4PD71 USYMA
66	16	57.1	499	2	Q6S60 ORYSA
67	16	57.1	504	2	Q9XZ8 LEIMA
68	16	57.1	534	2	Q4LZH5 9BURYK
69	16	57.1	538	2	Q86XHO HUMAN
70	16	57.1	539	1	GSHB ARATH
71	16	57.1	543	2	Q4QEB1 LEIMA
72	16	57.1	544	2	Q9F353 STRCO
73	16	57.1	556	2	Q7SN76 ACYTE
74	16	57.1	569	2	Q7QXT3 GIALLA
75	16	57.1	622	2	Q5JVP6 HUMAN
76	16	57.1	667	1	TS11 GIALLA
77	16	57.1	717	2	Q97444 GIALLA
78	16	57.1	717	2	Q4SB79 TETNG
79	16	57.1	719	2	Q9U019 GIALLA
80	16	57.1	719	2	Q9U021 GIALLA
81	16	57.1	729	2	Q9KH34 9GAMM
82	16	57.1	760	2	Q7OU65 GIALLA
83	16	57.1	772	2	Q5XXR3 RAT
84	16	57.1	776	1	ARHG6 HUMAN
85	16	57.1	776	2	Q7Z3W1 HUMAN
86	16	57.1	776	2	Q5YQ66 HUMAN
87	16	57.1	799	2	Q4T7H8 TETNG
88	16	57.1	868	2	Q7XPB6 ORYSA
89	16	57.1	1041	1	CHS1 CRYNY
90	16	57.1	1046	2	Q4SDRO TETNG
91	16	57.1	1089	2	Q4QBW2 LEIMA
92	16	57.1	1107	2	Q5CV70 CRYPV
93	16	57.1	1136	2	Q6CBRS YARLI
94	16	57.1	1167	2	Q57ZRO 9RYP
95	16	57.1	1178	2	Q6MM93 ORYSA
96	16	57.1	1236	2	Q5XDB2 CRYNNE
97	16	57.1	1341	2	Q4HV91 GIBZE
98	16	57.1	1494	2	Q4RL73 TETNG
99	16	57.1	1519	2	Q4P4P1 USYMA
100	16	57.1	1909	2	Q4Q998 LEIMA
101	16	57.1	5072	2	Q6TD66 TOKCO
102	16	53.6	35	2	Q5VUT7 9COCO
103	16	53.6	36	2	Q5VUT3 9COCO
104	15	53.6	43	2	Q53N14 ORYSA

Q67V9 ORYSA	Q5K9J0 CRYNNE	Q5SK09 CRYNNE	Q5K919 CRYNNE	Q4OSJ9 GLECA	Q7ONB5 GIALLA	Q8LD53 ARATH	Q9F3J7 ARATH	Q16421 CAEL	Q966B6 CAEL	Q16424 CAEL	Q9XV90 CAEL	Q8CAZ5 MOUSE	Q7PRQ7 ANOGA	Q6ZUB7 HUMAN	Q8PFE6 XANAC	Q6ZQJ1 HUMAN	Q9F934 BIPAD	Q89DY5 BRAJA	Q5OLQ5 ORYSA	Q5ZBR6 ORYSA	Q8N4O3 HUMAN	Q91E43 9ALPH	Q7WBR6 BORBA	Q7WPO8 BORBR	Q8XQX5 RALSO	Q55057 CRYNNE	Q5ZPV8 ORYSA	Q7Y021 ORYSA	Q941V9 ORYSA	Q7XUJ0 ORYSA	Q8BP04 MOUSE	Q7S1H3 NEUCR	Q4PD71 USYMA	Q6S60 ORYSA	Q9XZ8 LEIMA	Q4LZH5 9BURYK	Q86XHO HUMAN	GSHB ARATH	Q4QEB1 LEIMA	Q9F353 STRCO	Q7SN76 ACYTE	Q7QXT3 GIALLA	Q5JVP6 HUMAN	TS11 GIALLA	Q97444 GIALLA	Q4SB79 TETNG	Q9U019 GIALLA	Q9U021 GIALLA	Q9KH34 9GAMM	Q7OU65 GIALLA	Q5XXR3 RAT	ARHG6 HUMAN	Q7Z3W1 HUMAN	Q5YQ66 HUMAN	Q4T7H8 TETNG	Q7XPB6 ORYSA	CHS1 CRYNY	Q4SDRO TETNG	Q4QBW2 LEIMA	Q5CV70 CRYPV	Q6CBRS YARLI	Q57ZRO 9RYP	Q6MM93 ORYSA	Q5XDB2 CRYNNE	Q4HV91 GIBZE	Q4RL73 TETNG	Q4P4P1 USYMA	Q4Q998 LEIMA	Q6TD66 TOKCO	Q5VUT7 9COCO	Q5VUT3 9COCO	Q53N14 ORYSA
-------------	---------------	---------------	---------------	--------------	---------------	--------------	--------------	-------------	-------------	-------------	-------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	---------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	-------------	-------------	---------------	--------------	------------	--------------	--------------	--------------	---------------	--------------	-------------	---------------	--------------	---------------	---------------	--------------	---------------	------------	-------------	--------------	--------------	--------------	--------------	------------	--------------	--------------	--------------	--------------	-------------	--------------	---------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------

105	15	53.6	57	2	094MN5_9CAUD	094MN5 bacterioph	178	15	53.6	227	2	09VA83_DROME	09VA83 drosophila
106	15	53.6	57	2	04MOU9_BACCE	04mu9 bacillus ce	179	15	53.6	229	2	072969_9FUNG	072969 fungal endo
107	15	53.6	57	2	04MOV1_BACCE	04MOV1 bacillus ce	180	15	53.6	229	2	072973_9FUNG	072973 fungal endo
108	15	53.6	72	2	05C4E3_SCHJA	05C4E3 schistosoma	181	15	53.6	229	2	0729K3_9FUNG	0729K3 fungal ento
109	15	53.6	79	2	04MOV6_BACCE	04MOV6 bacillus ce	182	15	53.6	229	2	0729K4_9FUNG	0729K4 fungal ento
110	15	53.6	86	2	09YV33_9HIV1	09YV33 human immun	183	15	53.6	230	2	056P86_9METZ	056P86 uncultured
111	15	53.6	87	2	0757M4_NEUCR	0757M4 neurospora	184	15	53.6	234	2	056P87_9FUNG	056P87 fungal sp.
112	15	53.6	87	2	04RQO2_TERNG	04RQO2 tetraodon n	185	15	53.6	236	2	056P94_9METZ	056P94 uncultured
113	15	53.6	94	2	061G47_DROME	061G47 drosophila	186	15	53.6	239	2	096UV6_ORHPL	096UV6 optioetoma
114	15	53.6	94	2	070PFL_9VIRU	070PFL spodiopetra	187	15	53.6	242	2	051T33_MAGGR	051T33 magnaporthe
115	15	53.6	98	2	070PFL_9VIRU	070PFL spodiopetra	188	15	53.6	242	2	065XAS_ORYSA	065XAS oryza sativ
116	15	53.6	99	2	06UTV5_CRJPE	06UTV5 crinipellis	189	15	53.6	245	2	08FBS9_ECOLB	08FBS9 escherichia
117	15	53.6	99	2	04SDO4_TERNG	04SDO4 tetraodon n	190	15	53.6	248	2	061HM9_CABER	061HM9 caenorhabdi
118	15	53.6	106	2	06TS31_PABBR	06TS31 paracoccidi	191	15	53.6	248	2	045450_CABER	045450 caenorhabdi
119	15	53.6	107	2	06HEN9_ORYSA	06HEN9 oryza sativ	192	15	53.6	249	2	0624V6_CABER	0624V6 caenorhabdi
120	15	53.6	117	2	06HSN3_ORYSA	06HSN3 oryza sativ	193	15	53.6	250	2	0624V4_CABER	0624V4 caenorhabdi
121	15	53.6	125	2	084NA4_9FABA	084NA4 dialium guti	194	15	53.6	252	2	061135_CABER	061135 caenorhabdi
122	15	53.6	128	2	04L1I8_9BURK	04L1I8 burkholderi	195	15	53.6	252	2	0624V5_CABER	0624V5 caenorhabdi
123	15	53.6	132	2	08R5P9_MOUSE	08R5P9 mus musculu	196	15	53.6	257	2	07P2D8_ANGCA	07P2D8 anopheles g
124	15	53.6	139	2	072A85_DESYH	072A85 desulfovibr	197	15	53.6	257	2	0676U4_9SOLA	0676U4 nicotiana a
125	15	53.6	141	2	0825Q9_STRAM	0825Q9 streptomyce	198	15	53.6	261	2	06FPP4_CANGA	06FPP4 candida gla
126	15	53.6	150	2	05Z6U8_ORYSA	05Z6U8 oryza sativ	199	15	53.6	264	2	04Q805_LEITMA	04Q805 leishmania
127	15	53.6	152	2	06R217_9PEZI	06R217 arthrinitum	200	15	53.6	264	2	05DQ89_9CORA	05DQ89 canine coro
128	15	53.6	153	2	063176_BURPS	063176 burkholderi	201	15	53.6	268	2	022798_ARATH	022798 arabidopsis
129	15	53.6	154	2	066R16_YERPS	066R16 yerinia ps	202	15	53.6	269	2	06ZFE5_ORYSA	06ZFE5 oryza sativ
130	15	53.6	154	2	08ZB15_YERPE	08ZB15 yerinia pe	203	15	53.6	273	2	041070_PEA	041070 pisum sativ
131	15	53.6	155	2	07YSR5_DROME	07YSR5 drosophila	204	15	53.6	281	2	05RX10_ANGCA	05RX10 anopheles g
132	15	53.6	155	2	06Y1L9_DROYA	06Y1L9 drosophila	205	15	53.6	281	2	07RKP8_PLAYO	07RKP8 plasmodium
133	15	53.6	155	2	0721J0_DROSI	0721J0 drosophila	206	15	53.6	282	2	04LH85_9BURK	04LH85 burkholderi
134	15	53.6	155	2	0721J0_DROSI	0721J0 drosophila	207	15	53.6	282	2	08UFM9_AGRYS	08UFM9 agrobacteri
135	15	53.6	155	2	07XEA0_ORYSA	07XEA0 oryza sativ	208	15	53.6	285	2	05LXE3_SILPO	05LXE3 silicibacte
136	15	53.6	155	2	093TQ7_9DELT	093TQ7 uncultured	209	15	53.6	287	2	043528_LYCES	043528 lycopersico
137	15	53.6	156	2	0591P3_CANAL	0591P3 candida alb	210	15	53.6	289	2	043527_LYCES	043527 lycopersico
138	15	53.6	163	2	09B1M0_GIALA	09B1M0 giardia lam	211	15	53.6	293	2	061YT3_CABER	061YT3 caenorhabdi
139	15	53.6	163	2	07MQD6_VIBVY	07MQD6 vibrio vuln	212	15	53.6	294	2	09VTR3_DROME	09VTR3 drosophila
140	15	53.6	165	2	06R216_9PEZI	06R216 apiospora m	213	15	53.6	306	2	07F8S0_ORYSA	07F8S0 oryza sativ
141	15	53.6	165	2	08DDP4_VIBVU	08DDP4 vibrio vuln	214	15	53.6	318	2	07CZES_AGRYS	07CZES agrobacteri
142	15	53.6	169	2	04MHT3_BACCE	04MHT3 bacillus ce	215	15	53.6	332	2	062009_CERCA	062009 cercaria c
143	15	53.6	172	2	04NFB6_9MICA	04NFB6 arthropacte	216	15	53.6	334	2	08S202_DROME	08S202 drosophila
144	15	53.6	173	2	07UGY7_RHODA	07UGY7 rhodopirell	217	15	53.6	334	2	052V10_CIOIN	052V10 ciona intes
145	15	53.6	179	2	05ZZL9_MYCHY	05ZZL9 mycoplasma	218	15	53.6	335	2	09VW95_DROME	09VW95 drosophila
146	15	53.6	179	2	04NVV6_9DELT	04NVV6 anaeromyxob	219	15	53.6	335	2	052VH9_CIOIN	052VH9 ciona intes
147	15	53.6	181	2	061UD0_DROME	061UD0 drosophila	220	15	53.6	341	2	052V12_CIOIN	052V12 ciona intes
148	15	53.6	183	2	08S3W8_HORVD	08S3W8 hordeum vul	221	15	53.6	341	2	056N31_CIOIN	056N31 ciona intes
149	15	53.6	189	2	082P67_STRAM	082P67 streptomyce	222	15	53.6	342	2	053PA6_ORYSA	053PA6 oryza sativ
150	15	53.6	192	2	05SRJ2_CRYNE	05SRJ2 cryptococcu	223	15	53.6	343	2	091FHE_9INFA	091FHE influenza a
151	15	53.6	192	2	05RG00_CRYNE	05RG00 cryptococcu	224	15	53.6	343	2	04TFD4_TERNG	04TFD4 tetraodon n
152	15	53.6	193	2	07XIH6_ORYSA	07XIH6 oryza sativ	225	15	53.6	347	2	08SYW5_DROME	08SYW5 drosophila
153	15	53.6	195	2	070OG1_GIALA	070OG1 giardia lam	226	15	53.6	347	2	08G4A1_BIFIDA	08G4A1 bifidobacte
154	15	53.6	196	2	053M55_ORYSA	053M55 oryza sativ	227	15	53.6	350	2	06Z8S2_ORYDA	06Z8S2 oryza sativ
155	15	53.6	199	2	061G98_DROME	061G98 drosophila	228	15	53.6	352	2	0467B8_TERNG	0467B8 tetraodon n
156	15	53.6	209	2	04OG66_9LECA	04OG66 pertusaria	229	15	53.6	352	2	06P898_XENTR	06P898 xenopus tro
157	15	53.6	213	2	06V282_9BACT	06V282 symbiont ba	230	15	53.6	354	2	061LP0_DROME	061LP0 drosophila
158	15	53.6	214	2	06OQR7_CABER	06OQR7 caenorhabdi	231	15	53.6	357	2	09YV87_CABEL	09YV87 caenorhabdi
159	15	53.6	215	2	04OS18_9LECA	04OS18 pertusaria	232	15	53.6	362	1	MESD1_HUMAN	09Hk6 homo sapien
160	15	53.6	217	2	070YZ9_GIALA	070YZ9 giardia lam	233	15	53.6	362	1	MESD1_MOUSE	09Hk6 mus musculu
161	15	53.6	218	1	CK5F2_RAT	09J115 rattus norv	234	15	53.6	362	1	07QJL0_ANGCA	07QJL0 anopheles g
162	15	53.6	220	2	071T10_CHICK	071T10 gallus gall	235	15	53.6	362	2	04V3U8_DROME	04V3U8 drosophila
163	15	53.6	224	2	05KC04_CRYNE	05KC04 cryptococcu	236	15	53.6	362	2	05B4V3_RAT	05B4V3 rattus norv
164	15	53.6	225	2	04QSH8_9LECA	04QSH8 pertusaria	237	15	53.6	362	2	0542D0_MOUSE	0542D0 mus musculu
165	15	53.6	225	2	055XU2_CRYNE	055XU2 cryptococcu	238	15	53.6	362	2	08CFLO_MOUSE	08CFLO mus musculu
166	15	53.6	226	1	C553_PARDE	P29967 paracoccus	239	15	53.6	371	2	0510P9_HUMAN	0510F9 homo sapien
167	15	53.6	226	2	051672_PARDE	051672 paracoccus	240	15	53.6	377	2	09X1Y6_ORYSA	09X1Y6 oryza sativ
168	15	53.6	227	2	04QSG3_9LECA	04QSG3 pertusaria	241	15	53.6	379	2	07SEPF_ASRGO	07SEPF ashbya gos
169	15	53.6	227	2	04QSM6_9ASCO	04QSM6 coccoetema	242	15	53.6	385	2	07WFC3_TOXGO	07WFC3 toxoplasma
170	15	53.6	227	2	04QSM6_9ASCO	04QSM6 coccoetema	243	15	53.6	386	2	09VEY9_DROME	09VEY9 drosophila
171	15	53.6	227	2	04QSG8_9LECA	04QSG8 pertusaria	244	15	53.6	387	2	08KO10_BURCE	08KO10 burkholderi
172	15	53.6	227	2	04QSG8_9LECA	04QSG8 pertusaria	245	15	53.6	387	2	08KOJ7_BURCE	08KOJ7 burkholderi
173	15	53.6	227	2	04QSR6_9LECA	04QSR6 pertusaria	246	15	53.6	388	2	09B3Y6_ORHPO	09B3Y6 ophiocara p
174	15	53.6	227	2	04QSG0_9LECA	04QSG0 pertusaria	247	15	53.6	398	2	09B3Y2_9GOBI	09B3Y2 acanthogobi
175	15	53.6	227	2	04QSG16_9LECA	04QSG16 pertusaria	248	15	53.6	398	2	09B6C2_LATUA	09B6C2 lateolabrax
176	15	53.6	227	2	05JUE7_HUMAN	05JUE7 homo sapien	249	15	53.6	398	2	09BA00_9GOBI	09BA00 bostrychus
177	15	53.6	227	2	09UT7A3_DROME	09UT7A3 drosophila	250	15	53.6	398	2		

251	15	53.6	407	2	Q8SG80_MICSA	Q8sg80 micropterus	324	15	53.6	688	2	Q4PGR8_USTWA	Q4pgr8 ustilago ma
252	15	53.6	409	2	Q9SUA2_ARATH	Q9sua2 arabidopsis	325	15	53.6	694	2	Q4T212_TERNG	Q4t212 tetradon n
253	15	53.6	410	2	Q5HYM7_HUMAN	Q5hyt7 homo sapien	326	15	53.6	695	2	Q7XG03_ORYSA	Q7xg03 oryza sativ
254	15	53.6	411	1	CRPI_YARLI	P45815 yarrowia.1i	327	15	53.6	695	2	Q94HU6_ORYSA	Q94hu6 oryza sativ
255	15	53.6	415	2	Q5TVF2_ANOCA	Q5tvf2 anopheles g	328	15	53.6	705	2	Q7S136_NEUCR	Q7s136 neucora
256	15	53.6	416	2	Q6RK77_GIBZE	Q6rk77 gibberella	329	15	53.6	705	2	Q818W1_GIALA	Q818w1 giardia lam
257	15	53.6	419	1	Q5SRJ3_CRYNE	Q5srj3 cryptococcu	330	15	53.6	713	1	LRP10_HUMAN	Q724f1 homo sapien
258	15	53.6	420	1	Y461_STNY3	Q55167 synecocyst	331	15	53.6	713	2	Q5RD13_PONPY	Q5rd13 pongo pygma
259	15	53.6	422	2	Q7R075_GIALA	Q7r075 giardia lam	332	15	53.6	716	2	Q9UB84_9CILLI	Q9ub84 9cilli
260	15	53.6	423	2	Q9C8J4_ARATH	Q9c8j4 arabidopsis	333	15	53.6	727	2	Q8VL05_ALRISO	Q8vl05 aleromonas
261	15	53.6	428	2	Q4SEM6_TERNG	Q4sem6 tetradon n	334	15	53.6	731	2	Q8RT29_9GAMM	Q8rt29 pseudalter
262	15	53.6	433	2	Q9VMD4_DROME	Q9vmd4 drosophila	335	15	53.6	743	2	Q4T862_TERNG	Q4t862 tetradon n
263	15	53.6	442	2	Q5B5O6_EMENT	Q5b5o6 aspergillus	336	15	53.6	746	2	Q4SPR7_TERNG	Q4spr7 tetradon n
264	15	53.6	443	2	Q8LFE9_ARATH	Q8lfe9 arabidopsis	337	15	53.6	754	2	Q4SSB1_TERNG	Q4ssb1 tetradon n
265	15	53.6	444	2	Q93Z67_ARATH	Q93z67 arabidopsis	338	15	53.6	771	2	Q8K4I3_MOUSE	Q8k4i3 mus musculu
266	15	53.6	459	2	Q9XXP6_CABEL	Q9xxp6 caenorhabdi	339	15	53.6	772	2	Q4WK08_ASPFU	Q4wk08 aspergillus
267	15	53.6	459	2	Q4E207_LETMA	Q4e207 leishmania	340	15	53.6	799	2	Q5ONF9_ORYSA	Q5onf9 oryza sativ
268	15	53.6	460	2	Q725N8_DESYH	Q725n8 desulfocvibr	341	15	53.6	811	2	Q91582_HPBVO	Q91582 hepatitis b
269	15	53.6	461	2	Q58317_9TRYP	Q58317 trypanosoma	342	15	53.6	813	2	Q9BLJ2_CIOIN	Q9blj2 clona intes
270	15	53.6	464	2	Q7R682_GIALA	Q7r682 giardia lam	343	15	53.6	833	2	Q6J288_ACACA	Q6j288 acanthamoeb
271	15	53.6	465	2	Q4RETO_TERNG	Q4reto tetradon n	344	15	53.6	840	1	RNF19_MOUSE	P50636 mus musculu
272	15	53.6	470	2	Q8X0Z2_GIBRU	Q8x0z2 gibberella	345	15	53.6	843	2	Q91563_HPBVO	Q91563 hepatitis b
273	15	53.6	471	2	Q9VMG7_DROME	Q9vmg7 drosophila	346	15	53.6	843	2	Q91572_HPBVO	Q91572 hepatitis b
274	15	53.6	485	2	Q5U795_AZOLI	Q5u795 azospirillum	347	15	53.6	843	2	Q81141_HPBVO	Q81141 hepatitis b
275	15	53.6	489	1	ZIF1_CABEL	P34482 caenorhabdi	348	15	53.6	843	2	Q8UXG0_HPBVO	Q8uxg0 hepatitis b
276	15	53.6	490	2	Q7UPM3_RHOBA	Q7upm3 rhodospirill	349	15	53.6	869	2	Q8FGB3_ECOL6	Q8fgb3 escherichia
277	15	53.6	500	2	Q7UHJ6_RHOBA	Q7uhj6 rhodospirill	350	15	53.6	869	2	Q8PMF6_XANAC	Q8pmf6 xanthomonas
278	15	53.6	503	2	Q4YZM1_PLABE	Q4yzm1 plasmodium	351	15	53.6	885	2	Q6R389_9PROT	Q6r389 unclutred
279	15	53.6	505	2	Q4FKT0_9TRYP	Q4fkt0 trypanosoma	352	15	53.6	888	2	Q81PZ1_DROME	Q81pz1 drosophila
280	15	53.6	516	2	Q54C18_DICDI	Q54c18 dictyostell	353	15	53.6	904	2	Q4QAU0_LETMA	Q4qau0 leishmania
281	15	53.6	516	2	Q87P76_VIBPA	Q87p76 vibrio para	354	15	53.6	935	2	Q94719_PARTE	Q94719 parametium
282	15	53.6	520	2	Q8N1N4_HUMAN	Q8n1n4 homo sapien	355	15	53.6	937	2	Q5N1I1_FFRAT	Q5n1i1 francisella
283	15	53.6	521	2	Q7RTT2_HUMAN	Q7rtt2 homo sapien	356	15	53.6	943	2	Q6ODG4_ORYSA	Q6odg4 oryza sativ
284	15	53.6	522	2	Q5BAM2_EMENT	Q5bam2 aspergillus	357	15	53.6	948	2	Q5NXL5_AZOSE	Q5nxl5 azocarcus sp
285	15	53.6	525	2	Q5ZQI1_CRYNE	Q5zqi1 cryptococcu	358	15	53.6	969	2	Q8BR86_MOUSE	Q8br86 mus musculu
286	15	53.6	525	2	Q5KP13_CRYNE	Q5kpl3 cryptococcu	359	15	53.6	984	2	Q81PZ2_DROME	Q81pz2 drosophila
287	15	53.6	527	2	Q6ETD6_ORYSA	Q6etd6 oryza sativ	360	15	53.6	988	2	Q7QYU1_GIALA	Q7qyu1 giardia lam
288	15	53.6	531	1	PGLR1_RALSO	P58598 ralatonia s	361	15	53.6	993	2	Q9VOL8_DROME	Q9vol8 drosophila
289	15	53.6	538	2	Q51RL7_MAGRO	Q51rl7 magnaporthe	362	15	53.6	1002	1	SPHR_AMEPV	P29815 amastiza moo
290	15	53.6	541	2	Q4WTA0_ASPFU	Q4wta0 aspergillus	363	15	53.6	1003	2	Q5AMZ7_DICDI	Q5amz7 dictyostell
291	15	53.6	544	2	Q941V2_ORYSA	Q941v2 oryza sativ	364	15	53.6	1013	2	Q4SR96_TERNG	Q4sr96 tetradon n
292	15	53.6	549	2	P77818_AGRVI	P77818 agrobacteri	365	15	53.6	1023	2	Q5UB66_HUMAN	Q5ub66 homo sapien
293	15	53.6	563	2	Q6V9W0_RAT	Q6v9w0 rattus norv	366	15	53.6	1025	2	Q7R6J7_GIALA	Q7r6j7 giardia lam
294	15	53.6	564	2	Q9YQZ3_9HERP	Q9yqz3 randt herpe	367	15	53.6	1025	2	Q94EN9_CHLRE	Q94en9 chlamydomon
295	15	53.6	565	1	STK3_MOUSE	Q9zqg2 mus musculu	368	15	53.6	1039	2	Q610M2_CABER	Q610m2 caenorhabdi
296	15	53.6	569	2	Q9CDB3_MYCLE	Q9cdb3 mycobacteri	369	15	53.6	1043	2	Q17644_CABEL	Q17644 leishmania
297	15	53.6	574	2	Q8MXW1_HALKO	Q8mxw1 halocynthia	370	15	53.6	1057	2	Q4N4P8_THERA	Q4n4p8 thelletia p
298	15	53.6	587	2	Q9VHU6_DROME	Q9vhu6 drosophila	371	15	53.6	1061	2	Q4XQ34_PLACH	Q4xq34 plasmodium
299	15	53.6	589	2	Q4P936_USTWA	Q4p936 ustilago ma	372	15	53.6	1063	2	Q7QU10_GIALA	Q7qu10 giardia lam
300	15	53.6	593	2	Q818V8_GIALA	Q818v8 giardia lam	373	15	53.6	1068	2	Q4Q896_LETMA	Q4q896 leishmania
301	15	53.6	593	2	Q8LMC6_ORYSA	Q8lmc6 oryza sativ	374	15	53.6	1070	2	Q7R2W4_GIALA	Q7r2w4 giardia lam
302	15	53.6	595	2	Q5VXL7_HUMAN	Q5vxl7 homo sapien	375	15	53.6	1089	2	Q4Q9C2_LETMA	Q4q9c2 leishmania
303	15	53.6	599	2	Q8W4D7_ARATH	Q8w4d7 arabidopsis	376	15	53.6	1103	2	Q8WR15_DROME	Q8wr15 drosophila
304	15	53.6	602	2	Q6MG12_NEUCR	Q6mg12 neurospora	377	15	53.6	1146	2	Q5U4V9_XENIA	Q5u4v9 xenopus lae
305	15	53.6	605	2	Q9SFU8_ARATH	Q9sfu8 arabidopsis	378	15	53.6	1173	2	Q4T1A3_BACTS	Q4t1a3 bacillus th
306	15	53.6	608	2	Q7SDR2_NEUCR	Q7sdr2 neurospora	379	15	53.6	1186	1	C14AA_DACTS	Q4nuuv7 anaeromoxb
307	15	53.6	609	2	Q64BX6_9ARCH	Q64bx6 unclutred	380	15	53.6	1189	2	Q4NUU7_9DELT	Q4nuu7 pseudomonas
308	15	53.6	614	2	Q64DG6_9ARCH	Q64dg6 unclutred	381	15	53.6	1191	2	Q8RL60_PSEPT	Q8rl60 pseudomonas
309	15	53.6	619	2	Q8GZ21_ARATH	Q8gz21 arabidopsis	382	15	53.6	1216	2	Q4PDB3_USTWA	Q4pdb3 ustilago ma
310	15	53.6	622	2	Q8C9V4_MOUSE	Q8c9v4 mus musculu	383	15	53.6	1232	2	Q4T1X3_TERNG	Q4t1x3 tetradon n
311	15	53.6	623	2	Q818W0_GIALA	Q818w0 giardia lam	384	15	53.6	1236	2	Q80TK0_MOUSE	Q80tk0 mus musculu
312	15	53.6	635	2	Q7OP07_GIALA	Q7op07 giardia lam	385	15	53.6	1249	2	Q5TVB8_9TRYP	Q5tvb8 trypanosoma
313	15	53.6	637	2	Q4UIV3_THERA	Q4uiv3 thelletia a	386	15	53.6	1253	2	Q4T0S1_TERNG	Q4t0s1 tetradon n
314	15	53.6	638	2	Q7QOC4_GIALA	Q7qoc4 giardia lam	387	15	53.6	1285	2	Q8DA47_VIBVU	Q8da47 vibrio vuln
315	15	53.6	639	2	Q8PMW5_METWA	Q8pmw5 methanosarc	388	15	53.6	1297	2	Q7MK32_VIBVY	Q7mk32 vibrio vuln
316	15	53.6	639	2	Q8ITZ6_METAC	Q8itz6 methanosarc	389	15	53.6	1324	2	Q8SYK2_DROME	Q8syk2 drosophila
317	15	53.6	640	2	Q8C121_MOUSE	Q8c121 mus musculu	390	15	53.6	1324	2	Q9VTR8_DROME	Q9vtr8 drosophila
318	15	53.6	644	2	Q7OWT5_GIALA	Q7owt5 giardia lam	391	15	53.6	1343	2	Q4IK63_GIBZE	Q4ik63 gibberella
319	15	53.6	659	2	Q6C9W2_YARLI	Q6c9w2 yarrowia.1i	392	15	53.6	1353	2	Q6R2R5_ARATH	Q6r2r5 arabidopsis
320	15	53.6	659	2	Q4N6C1_THERA	Q4n6c1 thelletia p	393	15	53.6	1361	2	Q9NGV2_DROME	Q9ngv2 drosophila
321	15	53.6	669	2	Q4MMZ9_ASPFU	Q4mmz9 aspergillus	394	15	53.6	1361	2	Q9V7T4_DROME	Q9v7t4 drosophila
322	15	53.6	679	2	Q6PUC7_ANOCA	Q6puc7 anopheles g	395	15	53.6	1364	2	Q9SW45_ARATH	Q9sw45 arabidopsis
323	15	53.6	688	2	Q96VF2_USTWA	Q96vf2 ustilago ma	396	15	53.6	1376	2	Q8X1P2_PODAN	Q8x1p2 podospira a

397	15	53.6	1414	2	040545_LEIMA	04q645	leihamania	470	15	53.6	10495	2	04RE92_TETNG	04re92	tetradon n
398	15	53.6	1429	2	07F0W3_ANOGA	07p043	anopheles g	471	15	53.6	34350	2	08WZ42_HUMAN	08wz42	homo sapien
399	15	53.6	1447	2	04WPJ3_ASPEU	04wpj3	aspergillus	472	14	50.0		2	07W276_GASCO	07w276	gastricloria
400	15	53.6	1451	1	EM30_ARATH	042510	arabidopsis	473	14	50.0	21	2	05ZD63_MAGAR	05zd63	magnaporthe
401	15	53.6	1468	2	080TF6_MOUSE	080tf6	mus musculus	474	14	50.0	29	2	04TIR3_TETNG	04tir3	tetradon n
402	15	53.6	1514	2	04FX85_LEIMA	04fx85	leihamania	475	14	50.0	29	2	05YTU2_ANOGA	05ytu2	anopheles g
403	15	53.6	1516	2	07RLJ5_PLAYO	07rlj5	plasmodium	476	14	50.0	46	2	07RXU9_NEUCR	07rxu9	neurospora
404	15	53.6	1533	2	07RILO_PLAYO	07ri10	plasmodium	477	14	50.0	52	2	07JFK9_RHOBA	07jfk9	rhodospirella
405	15	53.6	1533	2	04Q5Q2_LEIMA	04q5q2	leihamania	478	14	50.0	53	2	07RSS9_RAT	07rss9	rattus norv
406	15	53.6	1539	2	07RT03_PLAYO	07rt03	plasmodium	479	14	50.0	55	2	05VIZ4_AERHY	05viz4	aeromonas h
407	15	53.6	1624	2	07RM69_PLAYO	07rm69	plasmodium	480	14	50.0	57	2	04TIC2_TETNG	04tic2	tetradon n
408	15	53.6	1625	2	04T364_TETNG	04t364	tetradon n	481	14	50.0	58	2	05EY69_ADE07	05ey69	human adeno
409	15	53.6	1651	2	07RNX5_PLAYO	07rx5	plasmodium	482	14	50.0	64	2	04YVM6_PLABE	04yvm6	plasmodium
410	15	53.6	1722	2	04NOK1_GDELT	04nkt1	anacronyxob	483	14	50.0	70	2	04YBP5_PLABE	04ybp5	plasmodium
411	15	53.6	1729	2	07XGP6_ORYSA	07xgp6	oryza sativ	484	14	50.0	72	2	084Z47_ORYSA	084z47	oryza sativ
412	15	53.6	1739	2	08ILZ0_ORYSA	08ilz0	oryza sativ	485	14	50.0	73	2	08RV11_PINPS	08rv11	pinus pins
413	15	53.6	1755	2	06RKI6_BOTCI	06rk16	botrytis ci	486	14	50.0	74	2	087AK8_XYLPT	087ak8	xyella tas
414	15	53.6	1782	2	04MQZ5_ASPEU	04mqz5	aspergillus	487	14	50.0	76	2	05TYD0_ANOGA	05tyd0	anopheles g
415	15	53.6	1782	2	06X012_SOLIN	06x012	solenopsis	488	14	50.0	77	2	089W35_BRAJA	089w35	bradyrhizob
416	15	53.6	1783	2	04FMM9_LEIMA	04fmm9	leihamania	489	14	50.0	77	2	09T023_9CAUD	09t023	bacterioph
417	15	53.6	1806	2	05BH30_EMENT	05bh30	aspergillus	490	14	50.0	78	2	07UHI8_RHOBA	07uhi8	rhodospirella
418	15	53.6	1833	2	04U0S1_BRABE	04u0s1	brachydanio	491	14	50.0	80	2	06TNI7_STWTH	06tni7	symblionace
419	15	53.6	1944	2	06MS81_9ACTO	06ms81	actinomadura	492	14	50.0	80	2	04T6X9_TETNG	04t6x9	tetradon n
420	15	53.6	1946	2	054JA3_DICDI	054ja3	dictyostell	493	14	50.0	82	2	04S1R2_TETNG	04s1r2	tetradon n
421	15	53.6	1958	2	084HN4_9ACTO	084hn4	kitasatospo	494	14	50.0	83	2	069MR0_ORYSA	069mr0	oryza sativ
422	15	53.6	1961	2	084HM9_9ACTO	084hm9	streptomyce	495	14	50.0	86	2	05P1S6_AZOGE	05p1s6	azococcus sp
423	15	53.6	1976	2	07PUM1_ANOGA	07pum1	anopheles g	496	14	50.0	88	2	097023_BREIN	097023	brevibacter
424	15	53.6	1977	2	08WU72_ANOGA	08wu72	anopheles g	497	14	50.0	88	2	09PPT1_XYLTA	09ppt1	xyella tas
425	15	53.6	1978	2	08TSH2_ANOGA	08tsh2	anopheles g	498	14	50.0	90	2	057D56_BRUVB	057d56	bruceella ab
426	15	53.6	1988	2	06RKI8_BOTCI	06rk18	botrytis ci	499	14	50.0	91	2	04XC49_PLACH	04xc49	placella ab
427	15	53.6	1996	2	04T5Z8_TETNG	04t5z8	tetradon n	500	14	50.0	91	2	091667_XENIA	091667	xenopus lae
428	15	53.6	2043	2	04Q510_LEIMA	04q510	leihamania	501	14	50.0	93	1	RS16_CAMP	RS16	campylob
429	15	53.6	2087	2	04S488_TETNG	04s488	tetradon n	502	14	50.0	93	2	0872F7_NEUCR	0872f7	neurospora
430	15	53.6	2088	2	P91443_CAEBL	P91443	caenorhabdi	503	14	50.0	93	2	06Z6F3_ORYSA	06z6f3	oryza sativ
431	15	53.6	2099	2	0622K8_CAEBR	0622k8	caenorhabdi	504	14	50.0	94	2	07PC59_MYCGM	07pc59	mycobacteri
432	15	53.6	2106	2	06RWD0_NECHA	06rwd0	neccria hae	505	14	50.0	97	2	P91216_CAEBL	P91216	caenorhabdi
433	15	53.6	2112	2	08WPLD_9UROC	08wpld	oikopleura	506	14	50.0	97	2	0531C1_9GAMA	0531c1	macaca fusc
434	15	53.6	2119	2	07H449_9GOST	07h449	geobacter s	507	14	50.0	98	2	07S7T5_NEUCR	07s7t5	neurospora
435	15	53.6	2159	2	09Y8A6_9PEZI	09y8a6	noduliporci	508	14	50.0	100	2	09YFR0_ABRPE	09yfr0	aeropyrum p
436	15	53.6	2159	2	086T05_DICDI	086t05	dictyostell	509	14	50.0	101	2	09DXR8_9CTAC	09dxr8	beak and fe
437	15	53.6	2159	2	0551N7_DICDI	0551n7	dictyostell	510	14	50.0	101	2	09DXG4_9CTAC	09dxg4	beak and fe
438	15	53.6	2162	2	08NR46_9PEZI	08nr46	xyella sp.	511	14	50.0	101	2	09DXG7_9CTAC	09dxg7	beak and fe
439	15	53.6	2162	2	09VOM0_DROME	09vom0	dirosophila	512	14	50.0	102	2	09WRM6_GRALE	09wrm6	gracillaria
440	15	53.6	2173	2	051P88_MAGGR	051p88	magnaporthe	513	14	50.0	102	2	0824W5_CHLGV	0824w5	chlamydomo
441	15	53.6	2176	2	046112_DROME	046112	dirosophila	514	14	50.0	103	1	Y040_MYCPA	Y040	mycoplasma
442	15	53.6	2181	1	STCA_EMENT	Q12397	emeritella	515	14	50.0	103	2	05YF63_9YVRU	05yf63	rock bream
443	15	53.6	2187	2	P79068_GLOIA	P79068	glomerella	516	14	50.0	104	2	013588_YEAST	013588	saccharomyc
444	15	53.6	2187	2	08TGD7_ASPEI	08tgd7	aspergillus	517	14	50.0	105	2	04TPW2_9EPHN	04tpw2	erythrobact
445	15	53.6	2188	2	06XR12_9PEZI	06xr12	ceratocysti	518	14	50.0	107	1	YB56_YEAST	YB56	yeast
446	15	53.6	2203	2	07S7J6_NEUCR	07s7j6	neurospora	519	14	50.0	107	2	005413_YEAST	005413	yeast
447	15	53.6	2238	2	05AVS5_EMENT	05avs5	aspergillus	520	14	50.0	108	2	007521_YEAST	007521	saccharomyc
448	15	53.6	2275	2	06RKE4_COCHR	06rke4	cochlioboln	521	14	50.0	109	2	06K452_ORYSA	06k452	oryza sativ
449	15	53.6	2397	2	06Q7Y4_PAPRR	06q7y4	paramectum	522	14	50.0	110	2	082YH0_STRAW	082yh0	strepomyce
450	15	53.6	2511	2	04ANK5_9DELT	04ank5	anaeromyxob	523	14	50.0	111	2	07SC68_NEUCR	07sc68	neurospora
451	15	53.6	2517	2	05AAZ33_EMENT	05aaz3	aspergillus	524	14	50.0	111	2	08IX35_HUMAN	08ix35	homo sapien
452	15	53.6	2533	2	P90589_PAPRR	P90589	paramectum	525	14	50.0	112	2	P72474_STRWU	P72474	strepococc
453	15	53.6	2533	2	027183_PAPRR	027183	paramectum	526	14	50.0	112	2	051906_DESVU	051906	desulfovibr
454	15	53.6	2536	2	P90649_PAPRR	P90649	paramectum	527	14	50.0	113	2	04V572_DROME	04v572	dirosophila
455	15	53.6	2556	2	07QJEB3_ANOGA	07qjeb3	anopheles g	528	14	50.0	113	2	09ZSE1_PINKA	09zse1	pinus radia
456	15	53.6	2644	2	06RKI2_BOTCI	06rk12	botrytis ci	529	14	50.0	113	2	056754_9HEPC	056754	hepatitis c
457	15	53.6	2703	1	NOTCH_DROME	P07207	dirosophila	530	14	50.0	115	2	0976E2_SULTO	0976e2	sulfolobus
458	15	53.6	2717	2	094710_PAPRR	094710	paramectum	531	14	50.0	116	2	08W633_9CAUD	08w633	bacterioph
459	15	53.6	2729	2	06P0K6_PAPRR	06p0k6	paramectum	532	14	50.0	116	2	0868W6_PLAPA	0868w6	plasmodium
460	15	53.6	2773	2	0591S5_BRABE	0591s5	brachydanio	533	14	50.0	118	2	08RPE3_FUSNN	08rpe3	fuobacteri
461	15	53.6	2793	2	05BBJ6_EMENT	05bbj6	aspergillus	534	14	50.0	119	2	04QAH0_LEIMA	04qah0	leihamania
462	15	53.6	2836	2	09V818_DROME	09v818	dirosophila	535	14	50.0	120	2	051901_DESBA	051901	desulfovibr
463	15	53.6	4039	2	04KCD6_PSEPS	04kcd6	pseudomonas	536	14	50.0	120	2	07P6A3_FUSNV	07p6a3	fusobacteri
464	15	53.6	4376	2	05MP07_9BACT	05mp07	symblion ba	537	14	50.0	121	2	084ZJ5_ORYSA	084zj5	oryza sativ
465	15	53.6	4767	2	017301_CAEBR	017301	caenorhabdi	538	14	50.0	121	2	04SZM9_TETNG	04szm9	tetradon n
466	15	53.6	4881	2	06TAB6_9BACT	06tab6	symblion ba	539	14	50.0	122	2	069HS6_CTOIN	069hs6	ciona intes
467	15	53.6	7743	2	0618F4_CAEBR	0618f4	caenorhabdi	540	14	50.0	123	2	06T9W1_NICSY	06t9w1	nicotiana s
468	15	53.6	7829	2	018559_CAEBL	018559	caenorhabdi	541	14	50.0	123	2	051904_9DELT	051904	desulfovibr
469	15	53.6	7962	2	Q10465_HUMAN	Q10465	homo sapien	542	14	50.0	123	2	08VXR8_MYCTU	08vxr8	mycobacteri

543	14	50.0	123	2	093464_CABAU	093464_carassius a	616	14	50.0	158	2	0641H1_9CIRC	0641h1 beak and fe
544	14	50.0	124	2	0759A8_NEUCR	0759A8_neupora	617	14	50.0	158	2	09DXF2_9CIRC	09dx2f beak and fe
545	14	50.0	124	2	08NB52_HUMAN	08NB52_homo sapien	618	14	50.0	158	2	09DXF5_9CIRC	09dx5f beak and fe
546	14	50.0	124	2	051916_9DELT	051916_unidentifie	619	14	50.0	159	2	06RZ20_9PEZI	06RZ20 periconia m
547	14	50.0	124	2	04VZP3_BURMA	04VZP3_burholderi	620	14	50.0	159	2	09PDM1_XYPIA	09PDM1 xyelia fas
548	14	50.0	125	2	06E1W5_9EURO	06E1W5_petromyces	621	14	50.0	160	2	04SYU6_TETNG	04SYU6 tetng
549	14	50.0	125	2	06E1W8_9EURO	06E1W8_aspergillus	622	14	50.0	161	2	06ZUT2_HUMAN	06ZUT2_homo sapien
550	14	50.0	125	2	06E1W9_9EURO	06E1W9_petromyces	623	14	50.0	161	2	095178_MACMA	095178 macaca mula
551	14	50.0	125	2	06E1X0_9EURO	06E1X0_petromyces	624	14	50.0	162	2	04IGT6_ADEP3	04IGT6 porcine ade
552	14	50.0	125	2	07R1V6_GIALA	07R1V6_giardia lam	625	14	50.0	162	2	04TH15_TETNG	04TH15 tetng
553	14	50.0	126	2	09JH18_MOUSE	09JH18_mus musculu	626	14	50.0	163	2	06RZ18_9EURO	06RZ18_talaromyces
554	14	50.0	129	1	LYSCI1_CANFA	LYSCI1_canis famli	627	14	50.0	163	2	09P2S0_HUMAN	09P2S0_homo sapien
555	14	50.0	129	1	05C456_SCHJA	05C456_schistosoma	628	14	50.0	163	2	0967R9_GIALA	0967R9_giardia lam
556	14	50.0	130	2	06RZ24_9PLEO	06RZ24_stemphylium	629	14	50.0	163	2	069X90_ORYSA	069X90_oryza sativ
557	14	50.0	131	1	NEU2_AANSN	PI9630_anser anser	630	14	50.0	164	2	06TGO1_9LECA	06TGO1 lecanora ma
558	14	50.0	132	2	07S1O7_NEUCR	07S1O7_neupospora	631	14	50.0	164	2	06TGO2_9LECA	06TGO2 lecanora fl
559	14	50.0	132	2	09V1J7_DROME	09V1J7_drosophila	632	14	50.0	164	2	06TGO3_9PEZI	06TGO3_9PEZI
560	14	50.0	132	2	08G4M6_BIFLO	08G4M6_bifidobacte	633	14	50.0	164	2	0458B6_CABEL	0458B6_caenorhabdi
561	14	50.0	133	1	YB9V_YEAST	P38350_saccharomyc	634	14	50.0	164	2	05SNE2_ORYSA	05SNE2_oryza sativ
562	14	50.0	133	2	0870B4_9PUNG	0870B4_pitomyces s	635	14	50.0	164	2	0445B7_CABEL	0445B7_caenorhabdi
563	14	50.0	133	2	05TVJ5_ANOGA	05TVJ5_anopheles g	636	14	50.0	165	2	09BHK2_9ANNE	09BHK2_sabella spa
564	14	50.0	134	2	08CTV1_HUMAN	08CTV1_homo sapien	637	14	50.0	166	2	05H5G2_XANOR	05H5G2_xanthomonas
565	14	50.0	134	2	06ZFU3_ORYSA	06ZFU3_oryza sativ	638	14	50.0	167	2	08YK95_MYCTU	08YK95_mycobacteri
566	14	50.0	135	2	05Q994_IXOSC	05Q994_ixodes scap	639	14	50.0	168	2	08H2J1_ORYSA	08H2J1_oryza sativ
567	14	50.0	135	2	018158_CABEL	018158_caenorhabdi	640	14	50.0	168	2	08REU4_9FLOR	08REU4_halymentia s
568	14	50.0	135	2	092P10_RHIME	092P10_rhizobium m	641	14	50.0	169	2	082G98_STRAW	082G98_streptomyces
569	14	50.0	136	2	061IL3_DROME	061IL3_drosophila	642	14	50.0	169	2	06ZMG5_HUMAN	06ZMG5_homo sapien
570	14	50.0	137	2	08TUJ9_MERAC	08TUJ9_methanosarc	643	14	50.0	170	2	07UZZ4_RHOBA	07UZZ4_rhodopirell
571	14	50.0	137	2	07Q4X7_ANOGA	07Q4X7_anopheles g	644	14	50.0	170	2	081NA4_ORYSA	081NA4_oryza sativ
572	14	50.0	137	2	08BWD7_MOUSE	08BWD7_m mus muscu	645	14	50.0	171	2	091B04_9ACTO	091B04_streptomyces
573	14	50.0	139	2	06RZ21_9EURO	06RZ21_dichotomomy	646	14	50.0	171	2	08SZV2_DROME	08SZV2_drosophila
574	14	50.0	139	2	000830_LEIMA	000830_leishmania	647	14	50.0	172	2	05B025_EWEMI	05B025_aspergillus
575	14	50.0	141	2	077076_HARCO	077076_haemonchus	648	14	50.0	174	2	06RZ25_9PLEO	06RZ25_westerdykel
576	14	50.0	141	2	084WF8_ORYSA	084WF8_oryza sativ	649	14	50.0	174	2	081PD4_DROME	081PD4_drosophila
577	14	50.0	143	1	PONA_DICTDI	P54660_dicyostell	650	14	50.0	174	2	05STU7_CRYNE	05STU7_cryptococcu
578	14	50.0	143	2	0585A5_9TRYP	0585A5_lycanosoma	651	14	50.0	175	1	SNAC_STRPR	SNAC_strpr
579	14	50.0	143	2	054BG4_DICTDI	054BG4_dicyostell	652	14	50.0	175	2	08GN71_BIFLO	08GN71_bifidobacte
580	14	50.0	143	2	06NA71_RHOBA	06NA71_rhodopseudo	653	14	50.0	175	2	05H2G2_XANOR	05H2G2_xanthomonas
581	14	50.0	145	2	07R3S4_GIALA	07R3S4_giardia lam	654	14	50.0	175	2	07WX94_ALCEU	07WX94_alcaligenes
582	14	50.0	148	2	052CL3_MAGGR	052CL3_magnaporthe	655	14	50.0	175	2	082PBE_STRAW	082PBE_streptomyces
583	14	50.0	148	2	08EVZ5_MYCPE	08EVZ5_mycoplasm	656	14	50.0	175	2	06RZ19_9ASCO	06RZ19_pseudonotr
584	14	50.0	149	2	06RZ22_9PEZI	06RZ22 cladospoiti	657	14	50.0	176	2	PHB8_POLBO	PHB8_porphyra te
585	14	50.0	149	2	07QXJ1_ANOGA	07QXJ1_anopheles g	658	14	50.0	177	1	PHB8_POLBO	PHB8_porphyra te
586	14	50.0	149	2	07XG18_ORYSA	07XG18_oryza sativ	659	14	50.0	177	1	PHB8_POLBO	PHB8_porphyra te
587	14	50.0	149	2	094I02_ORYSA	094I02_oryza sativ	660	14	50.0	177	1	PHB8_POLBO	PHB8_porphyra te
588	14	50.0	149	2	08AV79_BRARE	08AV79_brachydanio	661	14	50.0	177	1	PHB8_POLBO	PHB8_porphyra te
589	14	50.0	150	2	06RZ27_9EURO	06RZ27_pentacillium	662	14	50.0	177	1	PHB8_POLBO	PHB8_porphyra te
590	14	50.0	150	2	06TGP9_9LECA	06TGP9_lecanora st	663	14	50.0	177	2	09YD43_AERPE	09YD43_aeropyrum p
591	14	50.0	150	2	061641_ONCOC	061641_onchocerca	664	14	50.0	177	2	04WUY6_ASPEU	04WUY6_aspergillus
592	14	50.0	150	2	05NAT9_ORYSA	05NAT9_oryza sativ	665	14	50.0	177	2	08NAA9_HUMAN	08NAA9_homo sapien
593	14	50.0	151	2	07PKW7_ANOGA	07PKW7_anopheles g	666	14	50.0	177	2	09MRW8_GRALE	09MRW8_gracilaria
594	14	50.0	152	2	06RZ15_ASPEU	06RZ15_aspergillus	667	14	50.0	177	2	07S1P9_GRACH	07S1P9_gracilaria
595	14	50.0	154	2	09YAS6_AERPE	09YAS6_aeropyrum p	668	14	50.0	178	2	06RPM1_MOUSE	06RPM1_mus musculu
596	14	50.0	155	2	0857R5_9CAUD	0857R5_mycobacteri	669	14	50.0	178	2	09UAV9_CABEL	09UAV9_caenorhabdi
597	14	50.0	156	1	PPDA_ECOLI	P33554_escherichia	670	14	50.0	179	2	060WM3_CABER	060WM3_caenorhabdi
598	14	50.0	156	2	016422_CABEL	016422_caenorhabdi	671	14	50.0	179	2	067328_9INFA	067328_influenza a
599	14	50.0	156	2	06SEV4_9BACT	06SEV4_uncultured	672	14	50.0	181	2	07YTI7_CABEL	07YTI7_caenorhabdi
600	14	50.0	156	2	057KB8_SALCH	057KB8_salmonella	673	14	50.0	181	2	04LZX2_9BURK	04LZX2_burholderi
601	14	50.0	156	2	05PEM2_SALPA	05PEM2_salmonella	674	14	50.0	182	2	08CBW0_MOUSE	08CBW0_mus musculu
602	14	50.0	156	2	08ZMB0_SALTY	08ZMB0_salmonella	675	14	50.0	183	2	04T1Y1_CABEL	04T1Y1_caenorhabdi
603	14	50.0	156	2	082413_SALTY	082413_salmonella	676	14	50.0	184	2	069T53_ORYSA	069T53_oryza sativ
604	14	50.0	157	2	025068_HAEIR	025068_haemaphysa	677	14	50.0	184	2	08H447_ORYSA	08H447_oryza sativ
605	14	50.0	157	2	06XHV2_DROYA	06XHV2_drosophila	678	14	50.0	186	2	05TMH6_ANOGA	05TMH6_anopheles g
606	14	50.0	157	2	09BH65_GIALA	09BH65_giardia lam	679	14	50.0	188	1	DUS18_BOVIN	DUS18_bos taurus
607	14	50.0	157	2	09B1L6_GIALA	09B1L6_giardia lam	680	14	50.0	188	1	DUS18_HUMAN	DUS18_homo sapien
608	14	50.0	157	2	06AD65_LEIXX	06AD65_leishonia x	681	14	50.0	188	1	DUS18_MOUSE	DUS18_mus musculu
609	14	50.0	158	2	006775_HAEIR	006775_haemaphysa	682	14	50.0	188	1	DUS18_PONPY	DUS18_pongo pygma
610	14	50.0	158	2	09B1L8_GIALA	09B1L8_giardia lam	683	14	50.0	188	2	04R3T3_MACFA	04R3T3_macaca fasc
611	14	50.0	158	2	06SBR2_BACDI	06SBR2_bacillus li	684	14	50.0	189	2	0581F6_9CAUD	0581F6_cyanophora
612	14	50.0	158	2	090236_9CIRC	090236_bacillus li	685	14	50.0	189	2	09D9D8_MOUSE	09D9D8_mus musculu
613	14	50.0	158	2	0641F6_9CIRC	0641F6_beak and fe	686	14	50.0	189	2	052046_9ZZZZ	052046_plasmid phy
614	14	50.0	158	2	0641G0_9CIRC	0641G0_beak and fe	687	14	50.0	190	2	087TJ5_VIBPA	087TJ5_vibrio para
615	14	50.0	158	2	0641G7_9CIRC	0641G7_beak and fe	688	14	50.0	191	2	04WSA1_BACCE	04WSA1_bacillus ce

689	14	50.0	193	2	Q4PD06_USRTMA	Q4PD06_ustilago_ma	762	14	50.0	221	1	QAZ1_MESAU	P70112_mesocricetu
690	14	50.0	193	2	Q5TV08_ANOGA	Q5TV08_anopheles_g	763	14	50.0	221	2	Q4QSJ7_9LECA	Q4qsj7_pertuaria
691	14	50.0	193	2	Q5EM12_9COCO	Q5em12_avian_infec	764	14	50.0	221	2	Q8XEN5_CHUTE	Q8xens_chlorobium
692	14	50.0	195	2	Q54199_STREPCO	Q54199_streptomyce	765	14	50.0	221	2	Q5XJN5_BRABE	Q5xjns_brichydano
693	14	50.0	197	2	Q4WFM8_ASPFU	Q4wfm8_aspergillus	766	14	50.0	222	2	Q94722_PARTE	Q94722_paramecium
694	14	50.0	197	2	Q7R0J0_GIALA	Q7r0j0_giardia_lam	767	14	50.0	222	2	Q56PH2_RUEDE	Q56ph2_ruegeria_ge
695	14	50.0	197	2	Q44585_CAEEL	Q44585_caenorhabdi	768	14	50.0	224	2	Q56P81_9METZ	Q56p81_uncultured
696	14	50.0	197	2	Q8GGS6_STRAZ	Q8ggs6_streptomyce	769	14	50.0	225	2	Q20590_CABEIL	Q20590_caenorhabdi
697	14	50.0	198	1	DUS14_HUMAN	Q95147_homo_sapien	770	14	50.0	226	2	Q5PC56_SALPA	Q5pc56_salmonella
698	14	50.0	198	1	DUS14_MOUSE	Q911y7_homo_musculi	771	14	50.0	226	2	Q4RDM1_TETNG	Q4rdm1_tetradodon_n
699	14	50.0	198	2	Q6FI36_HUMAN	Q6fi36_homo_sapien	772	14	50.0	227	2	Q4QSG1_9LECA	Q4qsg1_pertuaria
700	14	50.0	198	2	Q8VKK4_MYCTU	Q8vkk4_mycobacteri	773	14	50.0	227	2	Q4QSI1_9LECA	Q4qsi1_pertuaria
701	14	50.0	198	2	Q8SV91_MOUSE	Q8sv91_mus_musculi	774	14	50.0	227	2	Q4QSI1_9LECA	Q4qsi1_pertuaria
702	14	50.0	198	2	Q7TPY1_MOUSE	Q7tpy1_mus_musculi	775	14	50.0	227	2	Q4QSN0_9ASCO	Q4qsn0_coccotrema
703	14	50.0	199	1	NIT1_EAST	P40472_saccharomyc	776	14	50.0	227	2	Q4QSN9_9ASCO	Q4qsn9_coccotrema
704	14	50.0	199	1	Q05729_YEAST	Q05729_saccharomyc	777	14	50.0	227	2	Q4QSM8_9ASCO	Q4qsm8_coccotrema
705	14	50.0	199	2	Q6QSA1_YEAST	Q6qsa1_saccharomyc	778	14	50.0	227	2	Q4QSM7_9ASCO	Q4qsm7_coccotrema
706	14	50.0	201	2	Q7PUG4_ANOGA	Q7pug4_anopheles_g	779	14	50.0	227	2	Q4QSM4_9LECA	Q4qsm4_pertuaria
707	14	50.0	201	2	Q6DGS6_BRARE	Q6dgs6_brachydanio	780	14	50.0	227	2	Q4QSM3_9LECA	Q4qsm3_octrolochia
708	14	50.0	202	2	Q4QSU5_9LECA	Q4qgs5_pertuaria	781	14	50.0	227	2	Q4QSU8_9LECA	Q4qsl8_octrolochia
709	14	50.0	202	2	Q4QIM4_LEIMA	Q4qim4_leishmania	782	14	50.0	227	2	Q4QSU5_9LECA	Q4qsl5_octrolochia
710	14	50.0	202	2	Q5H3M5_XANOR	Q5h3m5_xanthomonas	783	14	50.0	227	2	Q4QSL4_9LECA	Q4qsl4_pertuaria
711	14	50.0	203	2	Q4QEG4_LEIMA	Q4qeg4_leishmania	784	14	50.0	227	2	Q4QSL3_9LECA	Q4qsl3_pertuaria
712	14	50.0	203	2	Q4SFK7_TETNG	Q4sfk7_tetradodon_n	785	14	50.0	227	2	Q4QSL1_9LECA	Q4qsl1_pertuaria
713	14	50.0	204	1	DUS18_RAT	Q6axw7_rattus_novr	786	14	50.0	227	2	Q4QSK8_9LECA	Q4qsk8_pertuaria
714	14	50.0	204	2	Q5B220_EMENTI	Q5b220_aspergillus	787	14	50.0	227	2	Q4QSK5_9LECA	Q4qsk5_pertuaria
715	14	50.0	204	2	Q52472_NPYBM	Q92472_bomblyx_mori	788	14	50.0	227	2	Q4QSK4_9LECA	Q4qsk4_pertuaria
716	14	50.0	205	1	Y115_NPYOP	Q10154_oryzia_pseu	789	14	50.0	227	2	Q4QSK2_9LECA	Q4qsk2_pertuaria
717	14	50.0	205	2	Q94FA0_MIMGU	Q94fa0_mimulus_gut	790	14	50.0	227	2	Q4QSU6_9LECA	Q4qsj6_pertuaria
718	14	50.0	206	2	Q7Q606_ANOGA	Q7q606_anopheles_g	791	14	50.0	227	2	Q4QSU4_9LECA	Q4qsj4_pertuaria
719	14	50.0	206	2	Q8IHC4_DROME	Q8ihc4_drosophila	792	14	50.0	227	2	Q4QSU2_9LECA	Q4qsj2_pertuaria
720	14	50.0	206	2	Q8I9I4_DROME	Q8i9i4_drosophila	793	14	50.0	227	2	Q4QSI9_9LECA	Q4qsi9_pertuaria
721	14	50.0	206	2	Q6ZHF1_ORYSA	Q6zifi_oryza_sativ	794	14	50.0	227	2	Q4QSI5_9LECA	Q4qsi5_pertuaria
722	14	50.0	206	2	Q4SKY8_TETNG	Q4sky8_tetradodon_n	795	14	50.0	227	2	Q4QSI4_9LECA	Q4qsl4_pertuaria
723	14	50.0	206	2	Q89183_9RETR	Q89183_feline_immu	796	14	50.0	227	2	Q4QSI2_9LECA	Q4qsi2_pertuaria
724	14	50.0	207	2	Q4QSK0_9LECA	Q4qsk0_pertuaria	797	14	50.0	227	2	Q4QSI0_9LECA	Q4qsi0_pertuaria
725	14	50.0	209	2	Q94828_TETTH	Q94828_tetrahymena	798	14	50.0	227	2	Q4QSH9_9LECA	Q4qsh9_pertuaria
726	14	50.0	209	2	Q4NDI1_9MICC	Q4ndi1_archrobacte	799	14	50.0	227	2	Q4QSH7_9LECA	Q4qsh7_pertuaria
727	14	50.0	210	1	PRDX5_MOUSE	P99029_m_peroxired	800	14	50.0	227	2	Q4QSH3_9LECA	Q4qsh3_pertuaria
728	14	50.0	210	2	Q4QSH4_9LECA	Q4qsh4_pertuaria	801	14	50.0	227	2	Q4QSH2_9LECA	Q4qsh2_pertuaria
729	14	50.0	210	2	Q4X823_PLACH	Q4x823_plasmodium	802	14	50.0	227	2	Q4QSH0_9LECA	Q4qsh0_pertuaria
730	14	50.0	210	2	Q7TLR3_NPYCF	Q7tlr3_chorioconu	803	14	50.0	227	2	Q4QSH0_9LECA	Q4qsh0_pertuaria
731	14	50.0	211	2	Q4QSL7_9LECA	Q4qsl7_ochrolochia	804	14	50.0	227	2	Q4QSG7_9LECA	Q4qsg7_pertuaria
732	14	50.0	211	2	Q9D6X2_MOUSE	Q9d6x2_mus_musculi	805	14	50.0	227	2	Q4QSG5_9LECA	Q4qsg5_pertuaria
733	14	50.0	212	2	Q4QSM5_9ASCO	Q4qsm5_coccotrema	806	14	50.0	227	2	Q4QSG4_9LECA	Q4qsg4_pertuaria
734	14	50.0	212	2	Q4QSH5_9LECA	Q4qsh5_pertuaria	807	14	50.0	227	2	Q4QSG2_9LECA	Q4qsg2_pertuaria
735	14	50.0	212	2	Q5LA82_BACFN	Q5la82_bacteroides	808	14	50.0	227	2	Q4QSP8_9LECA	Q4qsf8_pertuaria
736	14	50.0	212	2	Q64OK1_BACFR	Q64ok1_bacteroides	809	14	50.0	227	2	Q4QSP4_9LECA	Q4qsf4_pertuaria
737	14	50.0	213	2	Q4QSK5_9LECA	Q4qsk5_pertuaria	810	14	50.0	227	2	Q4QSP3_9LECA	Q4qsf3_pertuaria
738	14	50.0	213	2	Q4QSH1_9LECA	Q4qsh1_pertuaria	811	14	50.0	227	2	Q4QSP1_9LECA	Q4qsf1_pertuaria
739	14	50.0	213	2	Q4QSP0_9LECA	Q4qsp0_pertuaria	812	14	50.0	227	2	Q4QSP8_9LECA	Q4qsf8_varicellari
740	14	50.0	213	2	Q4P5K8_USITMA	Q4p5k8_ustilago_ma	813	14	50.0	227	2	Q4QSP5_9LECA	Q4qsf5_pertuaria
741	14	50.0	213	2	Q54BI9_DICDI	Q54bi9_dicystosteli	814	14	50.0	227	2	Q4QSP5_9LECA	Q4qsf5_pertuaria
742	14	50.0	214	2	Q5L6N1_CHLAB	Q5l6n1_chlamydomoph	815	14	50.0	227	2	Q4QSL2_9LECA	Q4qsl2_pertuaria
743	14	50.0	214	2	Q9F6M0_RHJET	Q9f6m0_rhizobium_e	816	14	50.0	227	2	Q4QSK1_9LECA	Q4qsk1_pertuaria
744	14	50.0	214	2	Q53846_MYCTU	Q53846_mycobacteri	817	14	50.0	227	2	Q914U2_DROME	Q914u2_drosophila
745	14	50.0	214	2	Q7U158_MYCBO	Q7u158_mycobacteri	818	14	50.0	228	2	Q7JMT8_CABEIL	Q7jmt8_caenorhabdi
746	14	50.0	215	2	Q4WFS9_ASPFU	Q4wfs9_aspergillus	819	14	50.0	229	1	YAMB_RHISN	P55561_rhizobium_s
747	14	50.0	215	2	Q4PD16_USITMA	Q4pd16_ustilago_ma	820	14	50.0	229	2	Q7Z972_9PUNG	Q7z972_fungal_endo
748	14	50.0	216	2	Q51CC4_EMENTI	Q51cc4_emericeella	821	14	50.0	229	2	Q7Z9U8_9PUNG	Q7z9u8_fungal_endo
749	14	50.0	216	2	Q5AOG5_EBENI	Q5aog5_aspergillus	822	14	50.0	229	2	Q7Z9U9_9PUNG	Q7z9u9_fungal_endo
750	14	50.0	216	2	Q6KBI1_TREAVE	Q6kb19_tremetes_ve	823	14	50.0	229	2	Q7Z9K0_9PUNG	Q7z9k0_fungal_endo
751	14	50.0	216	2	Q4QSF2_9LECA	Q4qsf2_pertuaria	824	14	50.0	229	2	Q7Z9K1_9PUNG	Q7z9k1_fungal_endo
752	14	50.0	216	2	Q6ZM24_HUMAN	Q6zm24_homo_sapien	825	14	50.0	229	2	Q7Z9K2_9PUNG	Q7z9k2_fungal_endo
753	14	50.0	216	2	Q6NAG6_RHOPA	Q6nag6_rhodopseudo	826	14	50.0	229	2	Q51XB7_ORYSA	Q7b67_oryza_sativ
754	14	50.0	217	2	Q6V276_9BACT	Q6v276_symbiont_ba	827	14	50.0	230	2	Q51UX3_MAGRA	Q51ux3_magnaporthe
755	14	50.0	217	2	Q5P3R3_AZOSE	Q5p3r3_azaricus_sp	828	14	50.0	230	2	Q7Z970_9PUNG	Q7z970_fungal_endo
756	14	50.0	217	2	Q63PD8_BURPS	Q63pd8_burkholderi	829	14	50.0	231	2	Q8QSR7_9Z2Z2	Q8qsr7_fungal_endo
757	14	50.0	217	2	Q6ZEL4_BURMA	Q6zel4_burkholderi	830	14	50.0	231	2	Q8QSR7_9Z2Z2	Q8qsr7_fungal_endo
758	14	50.0	217	2	Q8BKP6_XANAC	Q8bkp6_xanthomonas	831	14	50.0	232	2	Q27423_DROVI	Q27423_drosophila
759	14	50.0	217	2	Q8QSR6_9Z2Z2	Q8qsr6_unidentifie	832	14	50.0	233	2	Q4RE76_TETNG	Q4re76_tetradodon_n
760	14	50.0	217	2	Q38020_9VIRU	Q38020_potato_viru	833	14	50.0	234	2	Q69HP3_CIOIN	Q69hp3_clona_inces
761	14	50.0	220	2	Q4QSH6_9LECA	Q4qsh6_pertuaria	834	14	50.0	234	2	Q9VIR4_DROME	Q9vir4_drosophila

835	14	50.0	234	2	Q98AP0_RHIL0	Q98AP0_rhizobium 1	908	14	50.0	270	2	Q46004_CABEL	Q46004_caenorhabdi
836	14	50.0	235	2	Q56PB1_9METZ	Q56PB1_uncultured	909	14	50.0	270	2	Q6EFC7_ACIAD	Q6EFC7_acinetobact
837	14	50.0	235	2	Q56PB3_9METZ	Q56PB3_uncultured	910	14	50.0	274	2	Q6UBK1_ZEADI	Q6UBK1_zea diplope
838	14	50.0	236	2	Q56P99_9METZ	Q56P99_uncultured	911	14	50.0	274	2	Q6TN86_METHX	Q6TN86_methylobact
839	14	50.0	236	2	Q4SZK8_TERNG	Q4SZK8_tetradodon n	912	14	50.0	274	2	Q9PPN6_XYLEA	Q9PPN6_xylella faa
840	14	50.0	239	2	Q9YB31_PASCO	Q9YB31_phoma sp. c	913	14	50.0	274	2	Q8TBC3_XYLET	Q8TBC3_xylella faa
841	14	50.0	239	2	Q5C403_SCHJA	Q5C403_schistosoma	914	14	50.0	274	2	Q8PLP2_XANAC	Q8PLP2_xanthomonas
842	14	50.0	239	2	Q7R376_GIALA	Q7R376_giardia lam	915	14	50.0	275	2	Q8L109_WHEAT	Q8L109_triticum ae
843	14	50.0	239	2	Q5BWM2_XANCA	Q5BWM2_xanthomonas	916	14	50.0	275	2	Q7W617_BORPA	Q7W617_bordeletia
844	14	50.0	239	2	Q5BWX0_XANCA	Q5BWX0_xanthomonas	917	14	50.0	275	2	Q7WIF9_BORPA	Q7WIF9_bordeletia
845	14	50.0	240	1	M810_AEATH	P92519_arabidopsis	918	14	50.0	275	2	Q7VYV2_BORPE	Q7VYV2_bordeletia
846	14	50.0	240	2	Q5IEN0_9ARCH	Q5IEN0_uncultured	919	14	50.0	276	2	Q96GV4_HUMAN	Q96GV4_homo sapien
847	14	50.0	240	2	Q5IEN3_9ARCH	Q5IEN3_uncultured	920	14	50.0	276	2	Q9NBD0_GLOMR	Q9NBD0_gloeosira mo
848	14	50.0	240	2	Q9Y864_ASPPA	Q9Y864_aspergillus	921	14	50.0	276	2	Q6UBK0_ZEADI	Q6UBK0_zea diplope
849	14	50.0	240	2	Q9HBS1_HUMAN	Q9HBS1_homo sapien	922	14	50.0	276	2	Q4UTS4_XANCP	Q4UTS4_xanthomonas
850	14	50.0	240	2	Q810D5_DROME	Q810D5_drosophila	923	14	50.0	276	2	Q5PEI1_SALPA	Q5PEI1_salmonella
851	14	50.0	241	2	Q9Y832_PENPA	Q9Y832_penicillium	924	14	50.0	276	2	Q8Z449_SALTI	Q8Z449_salmonella
852	14	50.0	241	2	Q9Y863_ASPPA	Q9Y863_aspergillus	925	14	50.0	276	2	Q8EPW2_XANCP	Q8EPW2_xanthomonas
853	14	50.0	242	2	Q5IEM9_9ARCH	Q5IEM9_uncultured	926	14	50.0	276	2	Q4S370_TERNG	Q4S370_tetradodon n
854	14	50.0	244	2	Q9Y100_9ARCH	Q9Y100_ectodeamia	927	14	50.0	277	2	Q9VFN8_DROME	Q9VFN8_drosophila
855	14	50.0	244	2	Q4U052_XANCP	Q4U052_xanthomonas	928	14	50.0	278	1	PLCB_MOUSE	Q8MK37_mus muscicu
856	14	50.0	244	2	Q4F6V7_9BACT	Q4F6V7_uncultured	929	14	50.0	278	2	Q8MQO2_DROME	Q8MQO2_drosophila
857	14	50.0	244	2	Q8PHR5_XANAC	Q8PHR5_xanthomonas	930	14	50.0	278	2	Q9ZTA3_RHIME	Q9ZTA3_rhizobium m
858	14	50.0	245	2	Q6RKX5_GIBZE	Q6RKX5_gibberella	931	14	50.0	280	2	Q8YWG3_LOLMU	Q8YWG3_lolium mult
859	14	50.0	248	2	Q8PAJ7_XANCP	Q8PAJ7_xanthomonas	932	14	50.0	281	2	Q8UKM4_VYIRU	Q8UKM4_helitotais z
860	14	50.0	249	2	Q5AU21_EMENI	Q5AU21_aspergillus	933	14	50.0	282	2	Q7XL74_ORYSA	Q7XL74_oryza sativ
861	14	50.0	249	2	Q60N68_CABER	Q60N68_caenorhabdi	934	14	50.0	283	2	Q7D973_MYCTU	Q7D973_mycobacteri
862	14	50.0	249	2	Q16420_CABEL	Q16420_caenorhabdi	935	14	50.0	285	1	C7A633_GIALA	Q7A633_giardia lam
863	14	50.0	250	1	BTG4_MOUSE	Q70552_mus muscicu	936	14	50.0	286	1	C7A633_GIALA	P28283_rattus norv
864	14	50.0	250	2	Q6I134_CABER	Q6I134_caenorhabdi	937	14	50.0	286	2	Q8BRN2_MOUSE	Q8BRN2_mus muscicu
865	14	50.0	250	2	Q16418_CABEL	Q16418_caenorhabdi	938	14	50.0	286	2	Q4RV58_TERNG	Q4RV58_tetradodon n
866	14	50.0	250	2	Q16419_CABEL	Q16419_caenorhabdi	939	14	50.0	287	2	Q8MWJ7_9ASCT	Q8MWJ7_bolletonia v1
867	14	50.0	250	2	Q5SND6_ORYSA	Q5SND6_oryza sativ	940	14	50.0	288	2	Q4I937_GIBBE	Q4I937_gibberella
868	14	50.0	250	2	Q5FUD7_GLIOX	Q5FUD7_glycomobact	941	14	50.0	288	2	Q9XV81_CABEL	Q9XV81_caenorhabdi
869	14	50.0	250	2	Q5HWV3_CAMJR	Q5HWV3_campylobact	942	14	50.0	289	2	Q6H604_ORYSA	Q6H604_oryza sativ
870	14	50.0	250	2	Q9ZST9_MOUSE	Q9ZST9_mus muscicu	943	14	50.0	289	2	Q7OV40_GIALA	Q7OV40_giardia lam
871	14	50.0	250	2	Q6NXZ4_MOUSE	Q6NXZ4_mus muscicu	944	14	50.0	289	2	Q9VU78_DROME	Q9VU78_drosophila
872	14	50.0	251	2	Q4P139_USTMA	Q4P139_ustilago ma	945	14	50.0	289	2	Q4V9U3_BRABE	Q4V9U3_briachydanto
873	14	50.0	253	2	Q4IACO_GIBZE	Q4IACO_gibberella	946	14	50.0	290	2	Q752M4_ASHGO	Q752M4_ashbya gos
874	14	50.0	253	2	Q6ZOX9_HUMAN	Q6ZOX9_homo sapien	947	14	50.0	290	2	Q9VX17_DROME	Q9VX17_drosophila
875	14	50.0	253	2	Q7UY70_RHOBA	Q7UY70_rhodopirell	948	14	50.0	290	2	Q6GHV3_BRABE	Q6GHV3_briachydanto
876	14	50.0	253	2	Q6UYU0_ORYSA	Q6UYU0_oryza sativ	949	14	50.0	291	2	Q7XZP8_ORYSA	Q7XZP8_oryza sativ
877	14	50.0	254	2	Q55700_PRRSV	Q55700_porcine rep	950	14	50.0	291	2	Q6H5L8_ORYSA	Q6H5L8_oryza sativ
878	14	50.0	254	2	Q513H3_PRRSV	Q513H3_porcine rep	951	14	50.0	291	2	Q6ES25_ORYSA	Q6ES25_oryza sativ
879	14	50.0	254	2	Q6A521_PRRSV	Q6A521_porcine rep	952	14	50.0	292	2	Q8RLY7_MOUSE	Q8RLY7_mus muscicu
880	14	50.0	254	2	Q84965_PRRSV	Q84965_porcine rep	953	14	50.0	293	2	Q7U3U8_SYNPX	Q7U3U8_synechococc
881	14	50.0	255	2	Q8N0W7_HUMAN	Q8N0W7_homo sapien	954	14	50.0	293	2	Q57124_9REOV	Q57124_coho salmon
882	14	50.0	255	2	Q56P90_9FUNG	Q56P90_fungal sp.	955	14	50.0	293	2	Q6ETW1_XENLA	Q6ETW1_xenopus lae
883	14	50.0	255	2	Q4Q1V5_LEIWA	Q4Q1V5_leishmania	956	14	50.0	296	2	Q4L1N2_GIBBE	Q4L1N2_gibberella
884	14	50.0	257	1	C1095_HUMAN	Q9H7C3_homo sapien	957	14	50.0	297	2	Q5ZNM2_9YIRU	Q5ZNM2_trypanosoma
885	14	50.0	257	2	Q5Z9G6_ORYSA	Q5Z9G6_oryza sativ	958	14	50.0	298	2	Q99159_TREBO	Q99159_cotesia con
886	14	50.0	258	2	Q60N66_CABER	Q60N66_caenorhabdi	959	14	50.0	298	2	Q5V869_ORYSA	Q5V869_oryza sativ
887	14	50.0	258	2	Q9Y0E9_DROME	Q9Y0E9_drosophila	960	14	50.0	298	2	Q7U7N8_SYNPX	Q7U7N8_synechococc
888	14	50.0	258	2	Q00946_9HYMN	Q00946_tetrahymena	961	14	50.0	299	2	Q5R0R8_IDILO	Q5R0R8_idiomarina
889	14	50.0	258	2	Q5PAG0_ANAMM	Q5PAG0_anaplasma m	962	14	50.0	300	2	Q2Z513_SOYBN	Q2Z513_glycine max
890	14	50.0	258	2	Q8QVQ2_ADABA	Q8QVQ2_bovine aden	963	14	50.0	300	2	Q4ZTF5_PRRSV	Q4ZTF5_pseudomonas
891	14	50.0	259	2	Q7S669_NEUCRA	Q7S669_neurospora	964	14	50.0	300	2	Q9YGD9_ONCMY	Q9YGD9_oncorhynch
892	14	50.0	259	2	Q83212_TREPA	Q83212_treponeema p	965	14	50.0	301	2	Q7V945_PROMM	Q7V945_prochloroc
893	14	50.0	261	2	Q94EL3_SORHL	Q94EL3_sorghum hal	966	14	50.0	301	2	Q8BN70_MOUSE	Q8BN70_mus muscicu
894	14	50.0	261	2	Q5BK83_RAT	Q5BK83_rattus norv	967	14	50.0	301	2	Q6ELZ9_BRABE	Q6ELZ9_briachydanto
895	14	50.0	262	2	Q5DC20_SCHJA	Q5DC20_schistosoma	968	14	50.0	301	2	Q4VBS7_BRABE	Q4VBS7_briachydanto
896	14	50.0	264	2	Q27253_CABEL	Q27253_mycobacteri	969	14	50.0	302	2	Q7X546_9ACTO	Q7X546_actinoplan
897	14	50.0	264	2	Q08343_MYCTU	Q08343_mycobacteri	970	14	50.0	302	2	Q4SZH4_TERNG	Q4SZH4_tetradodon n
898	14	50.0	264	2	Q7VFP02_MYGB0	Q7VFP02_mycobacteri	971	14	50.0	304	2	Q4LETA_9LILI	Q4LETA_sagittaria
899	14	50.0	267	2	Q6S5T5_FUSRU	Q6S5T5_fugu rubrip	972	14	50.0	304	2	Q9CUZ3_MOUSE	Q9CUZ3_mus muscicu
900	14	50.0	268	2	Q8I110_WHEAT	Q8I110_triticum ae	973	14	50.0	306	2	Q4UZB7_XANCP	Q4UZB7_xanthomonas
901	14	50.0	268	2	Q5H5W0_XANOR	Q5H5W0_xanthomonas	974	14	50.0	306	2	Q8PDI3_XANCP	Q8PDI3_xanthomonas
902	14	50.0	269	2	Q4Q5X9_LEIWA	Q4Q5X9_leishmania	975	14	50.0	307	2	Q94IDS_ORYSA	Q94IDS_oryza sativ
903	14	50.0	269	2	Q5GYN9_XANOR	Q5GYN9_xanthomonas	976	14	50.0	308	1	Q7R414_GIALA	Q7R414_giardia lam
904	14	50.0	269	2	Q6Z2Z2_BURPS	Q6Z2Z2_burkholderi	977	14	50.0	309	1	CF188_HUMAN	Q8NC51_homo sapien
905	14	50.0	269	2	Q8KKG2_BURPS	Q8KKG2_burkholderi	978	14	50.0	309	1	CF188_MOUSE	Q8F100_mus muscicu
906	14	50.0	269	2	Q8KKG3_BURPS	Q8KKG3_burkholderi	979	14	50.0	309	1	CF188_MOUSE	Q8F100_mus muscicu
907	14	50.0	269	2	Q6Z2U6_BURPS	Q6Z2U6_burkholderi	980	14	50.0	309	2	Q9VZG9_DROME	Q9VZG9_drosophila

```

RESULT 1
ID      SCS3 DROME      STANDARD;      PRT;      307 AA.
AC      P02840; O9VTU2;
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      10-MAY-2005 (Rel. 47, Last annotation update)
DE      Salivary glue protein Sgs-3 precursor.
GN      Name=Sgs3; ORFNames=CG11720;
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_TaxID=7227;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      MEDLINE=83294545; PubMed=6411930;
RT      Garkinkel M.D., Pruitt R.E., Meyerowitz E.M.;
RT      "DNA sequences, gene regulation and modular protein evolution in the
RT      Drosophila 68C glue gene cluster.";
RL      J. Mol. Biol. 168:765-789(1983).
[2]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=Berkely;
RC      MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RX      Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RX      Amaratunga P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RX      George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RX      Sutton G.G., Wortman J.R., Yandell M.D., Zhang L.X.,
RX      Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RX      Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RX      Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RX      Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RX      Beeson K.Y., Bencos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RX      Borovka D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RX      Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RX      Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RX      de Pablos B., Delcher A., Deng Z., Dey S., Dew I., Dietz S.M.,
RX      Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RX      Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RX      Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RX      Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RX      Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RX      Hoskins D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RX      Jalali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
RX      Jaiswal B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RX      Lasnik P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RX      Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

```

ALIGNMENTS

```

961      14      50.0      309      2      08P004_XANAC      08p004 xanthomaras
962      14      50.0      311      2      08V6X2_9VIRU      08v6x2 irs yellow
963      14      50.0      312      2      091PB3_9VIRU      091pb3 irs yellow
964      14      50.0      312      2      016417_CAEEL      016417 caenorhabd
965      14      50.0      315      2      070NT6_GIALA      070nt6 giardia lam
966      14      50.0      316      2      0680K2_ARATH      0680k2 arabidopsis
967      14      50.0      316      2      0817U1_ARATH      0817u1 arabidopsis
968      14      50.0      316      2      081AS1_ARATH      081as1 arabidopsis
969      14      50.0      317      2      06FHE1_HUMAN      06fhe1 homo sapien
990      14      50.0      317      2      08MWZ0_TRINI      08mwz0 trichoplusi
991      14      50.0      317      2      08BMY0_MOUSE      08bmy0 mus musculu
992      14      50.0      318      2      069AN2_9SAUR      069an2 rhinecort s
993      14      50.0      321      1      HEMW_WIGER      08d226 wig9leswort
994      14      50.0      321      1      07R109_GIALA      07r109 giardia lam
995      14      50.0      321      2      06NGE1_CORDI      06nge1 corynebacte
996      14      50.0      321      2      069AM6_9SAUR      069am6 xenochrophi
997      14      50.0      321      2      069AP5_9SAUR      069ap5 hemerophis
998      14      50.0      321      2      069AP7_9SAUR      069ap7 platycaps r
999      14      50.0      321      2      069AQ2_9SAUR      069aq2 platycaps r
1000     14      50.0      321      2      069AQ3_9SAUR      069aq3 platycaps r

```

```

RA      Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA      Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA      Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA      Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA      Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA      Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA      Spicer E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA      Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA      Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA      Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA      Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA      Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA      Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT      "The genome sequence of Drosophila melanogaster.";
RL      Science 287:2185-2195 (2000).
[3]
RP      GENOME REANNOTATION.
RX      MEDLINE=22426069; PubMed=12537572;
RX      Mista S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RX      Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.E.,
RX      Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RX      Belencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
RX      Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RX      Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RX      Lewis S.E.;
RT      "Annotation of the Drosophila melanogaster euchromatic genome: a
RT      systematic review.";
RL      Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
[4]
RP      NUCLEOTIDE SEQUENCE OF 1-28.
RX      MEDLINE=88332966; PubMed=3138416;
RT      Martin C.H., Mayeda C.A., Meyerowitz E.M.;
RT      "Evolution and expression of the Sgs-3 glue gene of Drosophila.";
RL      J. Mol. Biol. 201:273-287(1988).
[5]
RP      DEVELOPMENTAL STAGE.
RX      MEDLINE=94038699; PubMed=8223281;
RA      Hueb F., Ruiz C., Richards G.;
RT      "Puffs and PCR: the in vivo dynamics of early gene expression during
RT      ecdysone responses in Drosophila.";
RL      Development 118:613-627(1993).
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- TISSUE SPECIFICITY: Salivary gland specific.
CC      -1- DEVELOPMENTAL STAGE: In the salivary glands of mid instar larvae
CC      levels dramatically increase during puff stage 1 at 96-106 hours
CC      of development. Levels remain constant and abundant in late larvae
CC      until puff stage 10, then decrease by stage 11.
CC      -----
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      -----
DR      EMBL; X01918; CAA25994.1; -; Genomic DNA.
DR      EMBL; AE003544; AAF50056.1; -; Genomic DNA.
DR      EMBL; X78382; CAA55154.1; -; Genomic DNA.
DR      PIR; A03329; GSFF3.
DR      Ensembl; CG11720; Drosophila melanogaster.
DR      FlyBase; Fgn0003173; Sgs3.
KW      Repeat; signal.
FT      SIGNAL      1      23
FT      CHAIN      24      307      Potential,
FT                               salivary glue protein Sgs-3.
SQ      SEQUENCE      307 AA; 32196 MW; 45803DD1C418BC CR64;
OY      Query Match      64.3%; Score 18; DB 1; Length 307;
OY      Best local Similarity 16.7%; Pred. No. 4,9e-11;
OY      Matches      2; Conservative      0; Mismatches      10; Indels      0; Gaps      0;
OY      1 CXXXXXXXXXXC 12
OY      44 CTTTTTTTTTTC 55

```

```

RESULT 2
Q8K340_MOUSE PRELIMINARY; PRT; 347 AA.
ID Q8K340_MOUSE PRELIMINARY; PRT; 347 AA.
AC Q8K340_MOUSE PRELIMINARY; PRT; 347 AA.
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE 4632411J06Rik;
GN Name=4632411J06Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OC NCBI_Taxid=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CZECH II;
RC TISSUE=Mammary tumor metastasized to lung. MMTV-LTR/Mu1 model.
RX Expression driven by an MMTV-LTR enhancer.;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schlier G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stalcenon M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., Meewan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kerteman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield J.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CZECH II;
RC TISSUE=Mammary tumor metastasized to lung. MMTV-LTR/Mu1 model.
RC Expression driven by an MMTV-LTR enhancer.;
RA Strusberg R.;
RA Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
RA EMBL; BC028854; AAH28854.1; -; mRNA.
RA MGI; MGI:1925998; 4632411J06Rik.
RA GO; GO:0007165; P:signal transduction; IEA.
RA InterPro; IPR000159; RA.
RA PROSITE; PS50200; RA; 1.
FT NON_TER 1
SQ SEQUENCE 347 AA; 39161 MW; 306840D38A14B3DE CRC64;
Query Match 64.3%; Score 18; DB 2; Length 347;
Best Local Similarity 16.7%; Pred. No. 5.1e-11;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXX 12
DB 135 C6STSSSTASSC 146

```

```

DE enriched library, clone:6030474M07 product:hypothetical Serine-rich
DE region/Aminoacyl-transfer RNA synthetases class-II containing protein,
DE full insert sequence.
GN Name=4632411J06Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OC NCBI_Taxid=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka T.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsumoto Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Weng L.,
RA Sakai K., Oikawa T., Furuno M., Aono H., Balarelli R., Barsh G.,
RA Blake J., Boileau D., Bolunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gasterlstein S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaeerts P.,
RA Morone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Suganuma Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuura T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe K.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.

```


RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishii K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AK031660; BAC27498.1; -; mRNA.
 DR MGI; MGI:1925998; 463241J06Rik.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000159; RA.
 DR SMART; SM00314; RA; 1.
 DR PROSITE; PS50200; RA; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 362 AA; 40934 MW; 32B1047AC07B987 CRC64;

Query Match 64.3%; Score 18; DB 2; Length 362;
 Best Local Similarity 16.7%; Pred. No. 5.2e-11;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12
 DB 199 C8STSSSTSSC 210

RESULT 4
 O6AV38 ORYSA PRELIMINARY; PRT; 367 AA.
 AC O6AV38;
 DT 25-OCT-2004 (TRENBLrel. 28, Created)
 DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
 DE Putative Cyclin.
 GN Name=OSJNB0063J18.4;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 OX 1;
 RN NUCLEOTIDE SEQUENCE.
 RP Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
 RA Overton II L.L., Taitlin T., Kim M.M., Bera J.J., Jin S.S.,
 RA Padrosh D.W., Tallon L.J., Koo H., Ziemann V., Hsiao J., Blunt S.,
 RA Vanden S.S., Riedmuller S.B., Peterback T.T., Feldlyum T.V.,
 RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
 RA White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 3 BAC OSJNB0063J18 genomic sequence.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN NUCLEOTIDE SEQUENCE.
 RP Buell R.;
 RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the cyclin family.
 DR EMBL; AC107206; AAT77041.1; -; Genomic_DNA.
 DR Gramene; O6AV38; -;
 DR GO; GO:0005634; P:nucleus; IEA.
 DR GO; GO:000074; P:regulation of cell cycle; IEA.
 DR InterPro; IPR006670; Cyclin.
 DR InterPro; IPR004367; Cyclin_Cterm.
 DR InterPro; IPR006671; Cyclin_N.
 DR Pfam; PF02984; Cyclin_C; 1.
 DR Pfam; PF00134; Cyclin_N; 1.
 DR SMART; SM00385; CYCLIN; 1.
 KW Cyclin.
 SQ SEQUENCE 367 AA; 38727 MW; 023BF8C0CAD1D79 CRC64;

Query Match 64.3%; Score 18; DB 2; Length 367;
 Best Local Similarity 16.7%; Pred. No. 5.2e-11;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12
 DB 75 C8STSSSTSSC 86

RESULT 5
 O8BL43 MOUSE PRELIMINARY; PRT; 584 AA.
 ID O8BL43;
 AC O8BL43;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Mus musculus adult male corpora quadrigenina cDNA, RIKEN full-length
 DE enriched library, clone:B230384H21 product:hypothetical Serine-rich
 DE region/Aminoacyl-transfer RNA synthetases class-II containing protein,
 DE full insert sequence. (fragment).
 GN Name=463241J06Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=C57BL/6J; TISSUE=Corpora quadrigenina;
 RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RX Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."
 RL Meth. Enzymol. 303:19-44 (1999).
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=C57BL/6J; TISSUE=Corpora quadrigenina;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Koehli H.,
 RA Kuehl P., Lewis S., Mateno Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberte P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690 (2001).
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=C57BL/6J; TISSUE=Corpora quadrigenina;
 RC The PANTOM Consortium.
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=C57BL/6J; TISSUE=Corpora quadrigenina;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RX Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=20530913; PubMed=1076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Matshiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saio R., Saichou H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahita S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK046433; BAC32726.1; -; mRNA.
DR MGI: MGI:1925998; 4632411J06Rik.
DR GO: GO:0007165; P-signal transduction; IEA.
DR InterPro: IPR000159; RA.
DR SMART: SM00314; RA; 1.
DR PROSITE: PS50200; RA; 1.
KM Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 584 AA; 65345 MW; 941D03439114CC6E CRC64;
Query Match 64.3%; Score 18; DB 2; Length 584;
Best Local Similarity 16.7%; Pred. No. 5.9e-11;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 1 CXXXXXXXXXXC 12
Db 275 CSSTSSSTASC 286
RESULT 6
ID Q8LP13_PEA PRELIMINARY; PRT; 730 AA.
AC Q8LP13;
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE ABA insensitive 3.
GN Name=abi3;
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciales; Pisum.
OC NCBI_TaxID=3888;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Axillary bud;
RA Nakako A., Mori H.,
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB080195; BAC10553.1; -; mRNA.
DR GO: GO:0003677; F-DNA binding; IEA.
DR GO: GO:0006355; P-regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR003340; TF_B3.
DR Pfam: PF02362; B3; 1.

DR PROSITE: PS50863; B3; 1.
SQ SEQUENCE 730 AA; 80868 MW; 0A5BE9A7350A8B2 CRC64;
Query Match 64.3%; Score 18; DB 2; Length 730;
Best Local Similarity 16.7%; Pred. No. 6.2e-11;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 1 CXXXXXXXXXXC 12
Db 55 CSSSSSSASSSC 66
RESULT 7
ID Q9L012_ARATH PRELIMINARY; PRT; 35 AA.
AC Q9L012;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE F1504.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopses.
OC NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ecker J.R.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Cheuk R., Shin P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Alfieri H., Bel B., Chin C., Chou J., Choi E.,
RA Com L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharzky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thayeri A., Toriumi M., Vayberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC007887; AAF79343.1; -; Genomic DNA.
SQ SEQUENCE 35 AA; 3663 MW; CB4370A9D6A5D1F4 CRC64;
Query Match 60.7%; Score 17; DB 2; Length 35;
Best Local Similarity 16.7%; Pred. No. 1.9e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 1 CXXXXXXXXXXC 12
Db 15 CSIASTAAATSC 26
RESULT 8
ID Q24960_GIALA PRELIMINARY; PRT; 137 AA.
AC Q24960;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Trophozoite surface protein II (Fragment).
GN Name=TSPI1/L2;
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
OC NCBI_TaxID=5741;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98096860; PubMed=9435134;
RA Ey P.L., Mansouri M., Kaida J., Nohynkova E., Monis P.T.,
RA Andrews R.H., Mayrhofer G.,
RT "Genetic analysis of Giardia from hoofed farm animals reveals
RT antioxyt- specific and potentially zoonotic genotypes.";
RL J. Eukaryot. Microbiol. 44:626-635(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE.

```
RA Ey P.L., Mansouri M., Kulda J., Nohynkova E., Monis P.T.,  
RA Andrews R.H., Mayrhofer G.,  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U47631; AAB05399.1; -, Genomic_DNA.  
DR InterPro; IPR006209; EGF_1like.  
DR InterPro; IPR005127; Giardia_vsp.  
DR Pfam; PF03302; VSP; 1.  
DR PROSITE; PS01186; EGF_2; 1.  
FT NON_TER 137 137  
FT NON_TER 137 137  
SQ SEQUENCE 137 AA; 13875 MW; 42346969B5ECF37D CRC64;  
  
Query Match 60.7%; Score 17; DB 2; Length 137;  
Best Local Similarity 16.7%; Pred. No. 2.7e-09;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 22 CASATARAATC 33  
  
RESULT 9  
Q24987 GIALA PRELIMINARY; PRT; 137 AA.  
ID Q24987;  
AC Q24987;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Surface protein (Fragment).  
GN Name:stp;  
OS Giardia lamblia (Giardia intestinalis).  
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.  
OX NCBI_Taxid=5741;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Ey P.L., Darby J.M., Mayrhofer G.,  
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.  
DR EMBL; L16933; AAB59198.1; -, Genomic_DNA.  
DR InterPro; IPR005127; Giardia_vsp.  
DR Pfam; PF03302; VSP; 1.  
FT NON_TER 137 137  
FT NON_TER 137 137  
SQ SEQUENCE 137 AA; 14148 MW; A340072153005D58 CRC64;  
  
Query Match 60.7%; Score 17; DB 2; Length 137;  
Best Local Similarity 16.7%; Pred. No. 2.7e-09;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 22 CTSTARTAATC 33  
  
RESULT 10  
Q7XHS1 ORYZA PRELIMINARY; PRT; 210 AA.  
ID Q7XHS1;  
AC Q7XHS1;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE 2Fe-2S iron-sulfur cluster protein-like.  
GN Name:P0477A12.11;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhacridae; Oryzaceae; Oryza.  
OX NCBI_Taxid=39947;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Sasaki T., Matsumoto T., Katayose Y.,  
RT "Oryza sativa niponbare (GAS) genomic DNA, chromosome 7, PAC  
clone: P0477A12.";  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; AF005190; BAC80058.1; -, Genomic_DNA.  
DR Gramene; Q7XHS1; -;  
DR GO; GO:0005489; Electron transporter activity; IEA.  
DR GO; GO:0005506; F:iron ion binding; IEA.  
DR GO; GO:0046872; F:metal ion binding; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR01041; Ferredoxin.  
DR InterPro; IPR012675; Ferredoxin_fold.  
DR Pfam; PF01111; Fer2; 1.  
SQ SEQUENCE 210 AA; 22629 MW; 6BCDC4P408B21E1E CRC64;  
  
Query Match 60.7%; Score 17; DB 2; Length 210;  
Best Local Similarity 16.7%; Pred. No. 3e-09;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 25 CAVATATTSSC 36  
  
RESULT 11  
Q9XY90 GIALA PRELIMINARY; PRT; 234 AA.  
ID Q9XY90;  
AC Q9XY90;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Variant-specific surface protein (Fragment).  
GN Name=vsp417-4;  
OS Giardia lamblia (Giardia intestinalis).  
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.  
OX NCBI_Taxid=5741;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Ad-2;  
RX MEDLINE=99229863; PubMed=10215024; DOI=10.1016/S0166-6851(98)00183-2;  
RA Ey P.L., Darby J.M., Mayrhofer G.,  
RT "A new locus (vsp417-4) belonging to the tsa417-like subfamily of  
variant-specific surface protein genes in Giardia intestinalis.";  
RL Mol. Biochem. Parasitol. 99:55-68(1999).  
DR EMBL; AF065600; AAD28789.1; -, Genomic_DNA.  
DR InterPro; IPR006212; Furin_repeat.  
DR InterPro; IPR005127; Giardia_vsp.  
DR InterPro; IPR006210; IBGF.  
DR Pfam; PF03302; VSP; 1.  
DR SMART; SM00181; EGF; 2.  
DR SMART; SM00261; FU; 1.  
FT NON_TER 234 234  
FT NON_TER 234 234  
SQ SEQUENCE 234 AA; 23565 MW; FDC75E280AF7D517 CRC64;  
  
Query Match 60.7%; Score 17; DB 2; Length 234;  
Best Local Similarity 16.7%; Pred. No. 3.1e-09;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 22 CTSATARAATC 33  
  
RESULT 12  
O18039 CAEEL PRELIMINARY; PRT; 245 AA.  
ID O18039;  
AC O18039;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Hypothetical protein T06C12.14.  
GN ORFNames=T06C12.14;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI_Taxid=6239;
```

RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Brifcol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RG "The C. elegans sequencing consortium";
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology.";
 RL Science 282:2012-2018 (1998).
 DR EMBL: Z81116; CAB03300.1; -; Genomic_DNA.
 DR PIR: T24565; T24565.
 DR Ensembl: T06C12.14; Caenorhabditis elegans.
 DR Wormbase: WBGene00011521; T06C12.14.
 DR Wormpep: T06C12.14; CE16364.
 DR InterPro: IPR003582; ShKT.
 DR Pfam: PF01549; ShTK; 3.
 DR SMART: SM00254; ShKT; 4.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 245 AA; 25713 MW; BDB40649C966769A CRC64;
 Query Match 60.7%; Score 17; DB 2; Length 245;
 Best Local Similarity 16.7%; Pred. No. 3.1e-09;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 1 CXXXXXXXXXXC 12
 DB 193 CASATTTTSSCTC 204
 RESULT 13
 OS Q5QW08 CRYNE PRELIMINARY; PRT; 306 AA.
 AC Q5QW08;
 DT 25-OCT-2004 (TReMBLrel. 28, Created)
 DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
 DE Sec53p.
 GN Name=SEC53;
 OS Cryptococcus neoformans var. neoformans.
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
 OC Tremellomycetiales; Tremellales; Tremellaceae; Filobasidiella.
 OC NCB1_TaxID=40410;
 RX NCB1_TaxID=40410;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Jambon G.;
 RT "SEC53 encodes a putative phosphomannomutase.";
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY426175; AAR84595.1; -; Genomic_DNA.
 DR GO: GO:0005737; C:cytoplasm; IEA.
 DR GO: GO:0003824; F:catalytic activity; IEA.
 DR GO: GO:0004615; P:phosphomannomutase activity; IEA.
 DR GO: GO:0019307; P:mannose biosynthesis; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR006379; HAD_SF_IIB.
 DR InterPro: IPR005002; PMM.
 DR Pfam: PF03332; PMM; 1.
 DR TIGRFAMs: TIGR01484; HAD_SF_IIB; 1.
 SQ SEQUENCE 306 AA; 33844 MW; FB9DB2CF1A3EDDEA CRC64;
 Query Match 60.7%; Score 17; DB 2; Length 306;
 Best Local Similarity 16.7%; Pred. No. 3.3e-09;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 1 CXXXXXXXXXXC 12
 DB 5 CCGATTATTC 16
 RESULT 14
 OS Q6AS02 ORISA PRELIMINARY; PRT; 378 AA.
 AC Q6AS02;
 DT 25-OCT-2004 (TReMBLrel. 28, Created)
 DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
 DE Hypothetical protein OSUNBA0077J22.21.
 GN Name=OSUNBA0077J22.21;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoae; Oryza.
 OC NCB1_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
 RA Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
 RA Chao Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
 RA Heung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Liao H.-L.,
 RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-M., Yu C.-Y., Yu S.-W.,
 RA Wu H.-P., Shaw J.-F.,
 RT "Oryza sativa BAC OSUNBA0077J22 genomic sequence.";
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC148814; AAT93980.1; -; Genomic_DNA.
 DR Gramene; Q6AS02; -;
 KW Hypothetical protein.
 SQ SEQUENCE 378 AA; 37703 MW; 342586481D07BF5B CRC64;
 Query Match 60.7%; Score 17; DB 2; Length 378;
 Best Local Similarity 16.7%; Pred. No. 3.5e-09;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 1 CXXXXXXXXXXC 12
 DB 180 CVAAMAAATTAAC 191
 RESULT 15
 OS Q4LGT2_9BURK PRELIMINARY; PRT; 396 AA.
 ID Q4LGT2_9BURK
 AC Q4LGT2;
 DT 13-SEP-2005 (TReMBLrel. 31, Created)
 DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=Bcen242DRAFT_0168;
 OS Burkholderia cenocepacia H12424.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia; Burkholderia cecapia complex.
 OC NCB1_TaxID=331272;
 RX [1]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=H12424;
 RC US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Deter C., Glavina T.,
 RA Hammon N., Istrati S., Pitluck S., Richardson P.;
 RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia
 H12424.";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=H12424;
 RG US DOE Joint Genome Institute (JGI-ORNL);
 RA Larter P., Land M.;
 RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
 H12424.";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC EMBL: AALH0100139; EAM15270.1; -; Genomic_DNA.
 DR EMBL: AALH0100139;
 KW Hypothetical protein.
 SQ SEQUENCE 396 AA; 41189 MW; C829063DC312B057 CRC64;
 Query Match 60.7%; Score 17; DB 2; Length 396;
 Best Local Similarity 16.7%; Pred. No. 3.5e-09;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

QY      1 CXXXXXXXXXXC 12
      |
      18 CSSSDAAAAAAC 29

DB

RESULT 16
O8C8X1_MOUSE PRELIMINARY;      PRT;      429 AA.
ID      O8C8X1_MOUSE PRELIMINARY;      PRT;      429 AA.
AC      O8C8X1;
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Mus musculus adult retina cDNA, RIKEN full-length enriched library,
DE      clone:R330006D11 product:hypothetical protein, full insert
DE      Mus musculus (Mouse).
OS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC      Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=C57BL/6J; TISSUE=Retina;
RX      MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA      Carninci P., Hayashizaki Y.;
RT      "High-efficiency full-length cDNA cloning.";
RL      Meth. Enzymol. 303:19-44(1999).
RN      [2]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=C57BL/6J; TISSUE=Retina;
RX      MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA      Kawai T., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA      Arakawa T., Hara A., Fukunishi Y., Komoto H., Adachi T., Fukuda S.,
RA      Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA      Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA      Kadota K., Matsumoto H.A., Ashburner M., Bateman A., Casavant T.,
RA      Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA      Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA      Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA      Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA      Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA      Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA      Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA      Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbetsu P.,
RA      Norone P., Ring B., Ringwald M., Rodriguez I., Sakemoto N.,
RA      Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA      Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA      Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA      Hayashizaki Y.;
RT      "Functional annotation of a full-length mouse cDNA collection.";
RL      Nature 409:685-690(2001).
RN      [3]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=C57BL/6J; TISSUE=Retina;
RA      The PANTOM Consortium.
RA      the RIKEN Genome Exploration Research Group Phase I & II Team.
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs.";
RL      Nature 420:563-573(2002).
RN      [4]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=C57BL/6J; TISSUE=Retina;
RX      MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA      Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA      Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT      "Normalization and subtraction of cap-trapper-selected cDNAs to
RT      prepare full-length cDNA libraries for rapid discovery of new genes.";
RL      Genome Res. 10:1617-1630(2000).
RN      [5]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=C57BL/6J; TISSUE=Retina;
RX      MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA      Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,

```

```

RA      Kono H., Akiyama J., Nishi K., Katsunagi T., Tachiro H., Itoh M.,
RA      Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA      Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA      Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe K.,
RA      Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA      Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;
RT      "RIKEN integrated sequence analysis (RISA) system-384-format
RT      sequencing pipeline with 384 multiplexed sequencer.";
RL      Genome Res. 10:1757-1771(2000).
RN      [6]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=C57BL/6J; TISSUE=Retina;
RA      Adachi T., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA      Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA      Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA      Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA      Katoh H., Kawai J., Kojima Y., Kondo S., Komoto H., Koyama S.,
RA      Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA      Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA      Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA      Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA      Tagawa A., Takahashi F., Takaku-Akehira S., Takada Y., Tanaka T.,
RA      Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL      Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AK044299; BAC11860.1; -; mRNA.
DR      Ensembl; ENSMUSG00000014198; Mus musculus.
DR      GO; GO:0005634; C:nucleus; IEA.
DR      GO; GO:0003676; F:nucleic acid binding; IEA.
DR      GO; GO:0008270; F:zinc ion binding; IEA.
DR      InterPro; IPR007087; Znf_C2H2.
DR      InterPro; IPR003604; Znf_U1.
DR      SMART; SM00355; ZNF_C2H2_3.
DR      SMART; SM00451; ZNF_U1; 3.
DR      PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_3.
KW      Hypothetical protein.
SQ      SEQUENCE 429 AA; 44564 MW; 211C9F431P8F7FEF CRC64;

Query Match      60.7%; Score 17; DB 2; Length 429;
Best Local Similarity 16.7%; Pred. No. 3,6e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      1 CXXXXXXXXXXC 12
      |
      181 CDAASSSSSSC 192

DB

RESULT 17
O93H33_STRAW PRELIMINARY;      PRT;      940 AA.
ID      O93H33_STRAW PRELIMINARY;      PRT;      940 AA.
AC      O93H33;
DT      01-DEC-2001 (TrEMBLrel. 19, Created)
DT      01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT      13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE      Transport system protein (Putative oligopeptide ABC transporter
DE      substrate-binding protein).
GN      Name=oppA2; OrderedLocustNames=SAV1149;
OS      Streptomyces avermitilis.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Streptomycinae; Streptomycetaceae; Streptomycetes.
OX      NCBI_TaxID=33903;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX      MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.21143198;
RA      Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA      Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA      Kikuchi H., Shiba T., Sakai Y., Hattori M.;
RT      "Genome sequence of an industrial microorganism Streptomyces
RT      avermitilis: deducing the ability of producing secondary
RT      metabolites.";
RL      Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN      [2]
RP      NUCLEOTIDE SEQUENCE.

```

```

RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=2268306; PubMed=12692562; DOI=10.1038/bt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinohe M., Kikuchi H., Shiba T.,
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AB070952; BAB69358.1; -; Genomic DNA.
DR EMBL; BA000030; BAC70860.1; -; Genomic DNA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transporter; IEA.
DR InterPro; IPR000914; SBP_bac_5.
DR Pfam; PF00496; SBP_bac_5; 1.
SO Complete proteome.
QY SEQUENCE 540 AA; 58574 MW; 0052FEAD151B442C CRC64;

Query Match 60.7%; Score 17; DB 2; Length 540;
Best Local Similarity 16.7%; Pred. No. 3.8e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 12 CAAAAAATAATC 23

RESULT 18
ID Q9XTK3_GIALA PRELIMINARY; PRT; 667 AA.
AC Q9XTK3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE VSP417-3/A-II.
OS Name=vsp417-3/A-II;
GN Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BRIS/83/HEPU/136;
RX MEDLINE=99053029; PubMed=9836309;
RA By P.L., Darby J.M., Mayrhofer G.;
RT "Comparison of tsaa117-like variant-specific surface protein (VSP)
RT genes in Giardia intestinalis and identification of a novel locus in
RT genetic Group II isolates.";
RL Parasitology 117:445-455(1998).
DR EMBL; AF033584; AAP03497.1; -; Genomic DNA.
DR HSSP; OI6119; IEZG.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR005127; Giardia_VSP.
DR InterPro; IPR006210; IEGR.
DR Pfam; PF03302; VSP; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00261; FU; 4.
SO SEQUENCE 667 AA; 69123 MW; 77C64CFPS9441COC CRC64;

Query Match 60.7%; Score 17; DB 2; Length 667;
Best Local Similarity 16.7%; Pred. No. 4e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 448 CTSTARTATATC 459

RESULT 19
ID OSB3H3_EMENI PRELIMINARY; PRT; 1031 AA.
AC OSB3H3;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, last annotation update)
DE Hypoetical protein.

```

```

GN ORFNames=AN4907.2;
OS Aspergillus nidulans FGSC A4.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=227321;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FGSC A4;
RA Birren B., Nuebaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavsky L.,
RA Boukhalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Cornu B., DeArliano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hago B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kanai M., Kanat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Ton K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., McLean C., MacDonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Seery C., Rogov P.,
RA Roman J., Schauer S., Schuback R., Seaman S., Sivey P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Genome Sequence of Aspergillus nidulans.";
RT Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACD0100084; BAA60985.1; -; Genomic DNA.
KW Hypoetical protein.
SO SEQUENCE 1031 AA; 113626 MW; 6616443769CF2EF CRC64;

Query Match 60.7%; Score 17; DB 2; Length 1031;
Best Local Similarity 16.7%; Pred. No. 4.5e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 743 CTSSTSSASASTC 754

RESULT 20
ID MUC5B_HUMAN STANDARD; PRT; 5703 AA.
AC Q9HRC4; O00447; O00573; O14985; O15494; O95251; O95451; Q14881;
AC Q99552; Q9UE28;
DT 28-FEB-2003 (Ref. 41, Created)
DT 28-FEB-2003 (Ref. 41, last sequence update)
DT 13-SEP-2005 (Ref. 48, last annotation update)
DE Mucin-5B precursor (Mucin 5 subtype B, tracheobronchial) (High
DE molecular weight salivary mucin MG1) (Sublingual gland mucin).
GN Name=MUC5B; Synonyms=MUC5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE OF 1-1594.
RA Chen Y., Di Y.P., Wu R.;
RT "Molecular cloning of the amino-terminal and 5'-flanking region of the
RT human MUC5B mucin gene.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE OF 1-1325.
RX MEDLINE=99009274; PubMed=9790959; DOI=10.1006/dbrc.1998.9469;

```



```

FT CARBOHYD 5604 5604 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 5618 5618 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 5662 5662 N-linked (GlcNAc...) (Potential)
FT DISULFID 5594 5646 By similarity.
FT DISULFID 5622 5660 By similarity.
FT DISULFID 5626 5676 By similarity.
FT DISULFID 5643 5678 By similarity.
FT DISULFID 5645 5682 By similarity.
FT VARIANT 5137 5137 T->S (in dbSNP:2672788).
FT /FTID=VAR 014123.
FT CONFLICT 34 34 G->E (in Ref. 2).
FT CONFLICT 95 100 PDELGN->LPGLCK (in Ref. 2).
FT CONFLICT 104 104 S->C (in Ref. 2).
FT CONFLICT 142 142 E->K (in Ref. 1).
FT CONFLICT 225 225 R->S (in Ref. 2).
FT CONFLICT 330 331 PL->T (in Ref. 2).
FT CONFLICT 337 337 E->N (in Ref. 2).
FT CONFLICT 356 356 E->K (in Ref. 2).
FT CONFLICT 362 362 G->R (in Ref. 2).
FT CONFLICT 369 369 Missing (in Ref. 2 and 3).
FT CONFLICT 374 374 D->N (in Ref. 2).
FT CONFLICT 393 394 RT->TR (in Ref. 2).
FT CONFLICT 468 469 RK->GR (in Ref. 2).
FT CONFLICT 512 512 L->P (in Ref. 2).
FT CONFLICT 585 587 GAA->AH (in Ref. 3).
FT CONFLICT 601 601 A->S (in Ref. 3).
FT CONFLICT 628 629 DP->RS (in Ref. 2).
FT CONFLICT 633 633 F->L (in Ref. 2).
FT CONFLICT 676 676 A->P (in Ref. 3).
FT CONFLICT 701 701 R->P (in Ref. 3).
FT CONFLICT 752 752 E->K (in Ref. 2).

```

Query Match
Best Local Similarity 60.7%; Score 17; DB 1; Length 5703;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 5062 CTSASATMAARC 5073

```

RESULT 21
O6EUA6 ORYSA PRELIMINARY; PRT; 93 AA.
AC O6EUA6.
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DE Hypothetical protein OJ1116_C12.20.
GN Name=OJ1116_C12.20;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCB1_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RA Submitted (Sep-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004134; BAD27763.1; -; Genomic_DNA.
DR Gramene; O6EUA6; -.
KW Hypothetical protein.
SQ SEQUENCE 93 AA; 9912 MW; 10814C0A5B0F123 CRC64;

```

Query Match
Best Local Similarity 57.1%; Score 16; DB 2; Length 93;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 14 CTSASAGDARTSC 25

```

RESULT 22
O850L9 ORYSA PRELIMINARY; PRT; 100 AA.
AC O850L9.
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DE Putative topoisomerase.
GN Name=OSJNBA0032E21.11;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCB1_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Eastman A.P., Smith S.C., Bertin N., Liang C., Najjar F.Z., Pratt L.H.,
RA Cordomier-Pratt M.-M.;
RA Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF377947; AA032310.1; -; Genomic_DNA.
DR Gramene; O850L9; -.
DR GO; GO:0016853; P:isomerase activity; IEA.
KW Isomerase.
SQ SEQUENCE 100 AA; 11026 MW; FA5BF1C9B0965999 CRC64;

```

Query Match
Best Local Similarity 57.1%; Score 16; DB 2; Length 100;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 60 CSSASASAAAPC 71

```

RESULT 23
O62UJ1 BURMA PRELIMINARY; PRT; 100 AA.
AC O62UJ1.
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DE Hypothetical protein.
GN OrderedOcusaNames=BMA1578;
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCB1_TaxID=13573;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.040306101;
RA Nieman W.C., Deshazer D., Kim H.S., Tetteijn H., Nelson K.E.,
RA Feldlyum T.V., Ulrich R.L., Roming C.M., Brinkac L.M.,
RA Daugherty S.C., Davidson T.D., Deboy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouli H.M., Kolonay J.F.,
RA Madupu R., Mohammed Y., Nelson W.C., Kadane D., Romero C.M.,
RA Sarria S., Selengut J., Shambhlin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome."
RT Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251 (2004).
DR EMBL; CP000010; AAU47930.1; -; Genomic_DNA.
DR TIGR; BMA1578; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 100 AA; 11000 MW; 89664E4C442D9F5 CRC64;

```

Query Match
Best Local Similarity 57.1%; Score 16; DB 2; Length 100;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 20 CTAATAATTATC 31

```

RESULT 24
CBPI_AJBECA STANDARD; PRT; 110 AA.
ID CBPI_AJBECA
AC 042720;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last annotation update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Calcium-binding protein precursor (CBP).
GN Name=CBP1;
OS Ajellomyces capsulata (Histoplasma capsulatum).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Oxyenales; Ajellomycetaceae; Ajellomyces.
OX NCBI_TaxID=5037;
RN NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.
RP STRAIN=G186AR;
RC MEDLINE=98149309; PubMed=9489665;
RA Batnaghari J.W., Deepe G.S. Jr., Di Cara E., Goldman W.E.;
RT "Histoplasma acquisition of calcium and expression of CBP1 during
RT intracellular parasitism.";
RL Mol. Microbiol. 27:531-539 (1998).
RN CHARACTERIZATION, AND MASS SPECTROMETRY.
RP STRAIN=G186AR;
RC MEDLINE=98053973; PubMed=9393824;
RA Batnaghari J.W., Goldman W.E.;
RT "Calcium dependence and binding in cultures of Histoplasma
RT capsulatum.";
RL Infect. Immun. 65:5257-5261 (1997).
CC -!- FUNCTION: Involved in calcium binding and uptake in yeast phase.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- DEVELOPMENTAL STAGE: Expressed in the yeast phase during
CC infection, and not in mycelia.
CC -!- PTM: May be glycosylated.
CC -!- MASS SPECTROMETRY: MW=7858.0; METHOD=Electrospray; RANGE=33-110;
CC NOE=Ref.2.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF006209; AAC39354.1; -; Genomic_DNA.
CC KW Calcium; Direct protein sequencing; Glycoprotein; Signal.
CC FT SIGNAL 1 19 Potential.
CC FT CHAIN 1 110 Calcium-binding protein.
CC FT DISULFID 60 93 Potential.
CC FT DISULFID 65 76 Potential.
CC FT DISULFID 105 108 Potential.
CC SQ SEQUENCE 110 AA; 11014 MW; BD98F2FA1B4C77CE CRC64;

Query Match 57.1%; Score 16; DB 1; Length 110;
Best Local Similarity 16.7%; Pred. No. 1.7e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 65 CUSLSASSAAC 76

```

```

OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Han Z.;
RT Submitted (MAR-2005) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY810427; AAX26316.1; -; mRNA.
DR InterPro; IPR009081; ACP-like.
DR InterPro; IPR011989; ARM-like.
DR InterPro; IPR011992; EF-Hand_type.
KW Hypothetical protein.
SQ SEQUENCE 144 AA; 16762 MW; FF652713F976647B CRC64;

Query Match 57.1%; Score 16; DB 2; Length 144;
Best Local Similarity 16.7%; Pred. No. 1.8e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 15 CSSYAVSSTSC 26

RESULT 26
O9YBV0_AERPE PRELIMINARY; PRT; 152 AA.
ID O9YBV0_AERPE
AC O9YBV0;
DT 01-NOV-1999 (TRMBLrel. 12, Created)
DT 01-NOV-1999 (TRMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TRMBLrel. 24, Last annotation update)
DE Hypothetical protein APL1499.
DE OrderedAccession=APL1499;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KL;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobid hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix KL.";
RL DNA Res. 6:83-101 (1999).
DR EMBL; BA000002; BAA80498.1; -; Genomic_DNA.
DR PIR; D72630; D72630.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 152 AA; 16241 MW; 4386978A0D6151EF CRC64;

Query Match 57.1%; Score 16; DB 2; Length 152;
Best Local Similarity 16.7%; Pred. No. 1.8e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 137 CSSTATTLSTRC 148

RESULT 27
O7SLM9_ORYSA PRELIMINARY; PRT; 161 AA.
ID O7SLM9_ORYSA
AC O7SLM9;
DT 05-JUL-2004 (TRMBLrel. 27, Created)
DT 05-JUL-2004 (TRMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRMBLrel. 27, Last annotation update)
DE Putative DNA topoisomerase (with alternative splicing).
GN Name=OSJNBa0047E24.8;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```



```
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Euphorbiaceae; Oryzae; Oryza.
OX NCBI_Taxid=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Taitlin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadresh D.W., Tallon L.J., Koo H., Ziemann V., Heiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Uteback T.T., Feldlyum T.V.,
RA Yang Q.Q., Haas B.J., Sub B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RN Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RL [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RN Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC092556; AAR87258.1; -; Genomic_DNA.
DR Gramene; Q75LM9; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR InterPro; IPR002815; DNA_topII.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF04406; TPGA_N; 1.
DR PRINTS; PR01550; TOF6AFAMILY.
DR DNA-binding; Isomerase; Transcription.
SQ SEQUENCE 161 AA; 18051 MW; 41B12D72B71B76C3 CRC64;

Query Match
Best Local Similarity 57.1%; Score 16; DB 2; Length 161;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 60 CSSASAAAAPC 71

RESULT 28
Q7PY44_ANOGA PRELIMINARY; PRT; 172 AA.
ID Q7PY44;
AC Q7PY44;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE ENSANGP00000012250 (Fragment).
GN ORFNames=ENSANGG00000009761;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_Taxid=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=PEST;
RN Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA80100887; EAA01398.1; -; Genomic_DNA.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PR001010; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
FT NON_TER 1
SQ SEQUENCE 172 AA; 18464 MW; E857469A135A1884 CRC64;

Query Match
Best Local Similarity 57.1%; Score 16; DB 2; Length 172;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 CXXXXXXXXXXC 12
DB 77 CSSSSSSAGSEC 88

RESULT 29
Q6ZD90_ORYSA PRELIMINARY; PRT; 176 AA.
ID Q6ZD90;
AC Q6ZD90;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Hypothetical protein P0438H08.27.
GN Name=P0438H08.27;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Euphorbiaceae; Oryzae; Oryza.
OX NCBI_Taxid=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RN Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; AP004460; BAC99511.1; -; Genomic_DNA.
DR Gramene; Q6ZD90; -.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PR00097; zf-C3HC4_1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2.1.
DR Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 176 AA; 18608 MW; 79FE9D42C740112 CRC64;

Query Match
Best Local Similarity 57.1%; Score 16; DB 2; Length 176;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 96 CRRAAAAAASC 107

RESULT 30
Q55014_DICDI PRELIMINARY; PRT; 187 AA.
ID Q55014;
AC Q55014;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DDB0217903;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_Taxid=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugang R., Berrihan M., Song J., Olesch R., Szefranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rest R., Church C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Landay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
```

```
RA Louised H., Mungall K., Oliver K., Price C., Quail M.A.,  
RA Urushihara H., Hernandez J., Rabbittowitch E., Steffen D., Sanders M.,  
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivy A.,  
RA Sugano S., White B., Walker D., Woodward J., Winkler T., Tanaka Y.,  
RA Shalinsky G., Glibcher R., Weinstein G., Rosenhal A., Cox E.C.,  
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,  
RA Williams V., Dear P.H., Noegel A.A., Bartell B., Kuapa A.,  
RT "The genome of the social amoeba Dictyostelium discoideum.";  
RT Nature 0:0-0(2005).  
CC -1- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data  
CC EMBL:AAFI0100042; EMBL047.1; -; Genomic_DNA.  
DR EMBL:AAFI0100042; EMBL047.1; -; Genomic_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 187 AA; 20124 MW; 8D78DE804DFBC104 CRC64;  
  
Query Match 57.1%; Score 16; DB 2; Length 187;  
Best Local Similarity 16.7%; Pred. No. 1.9e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 65 CYSYSSSSSSSFC 76  
  
RESULT 31  
Q82LKO_STRAM PRELIMINARY; PRT; 208 AA.  
ID Q82LKO;  
AC Q82LKO;  
DT 01-JUN-2003 (TRMBLrel. 24, Created)  
DT 01-JUN-2003 (TRMBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TRMBLrel. 24, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocusNames=SAV2010;  
OS Streptomyces avermitilis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycetaceae; Streptomycetaceae; Streptomycetes;  
OC NCBI_Taxid=33903;  
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;  
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
RA Sakaki Y., Hattori M., Omura S.;  
RT "Complete genome sequence and comparative analysis of the industrial  
RT microorganism Streptomyces avermitilis.";  
RT Nat. Biotechnol. 21:526-531(2003).  
RN [2]  
RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).  
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
RC MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.21143198;  
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,  
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
RT "Genome sequence of an industrial microorganism Streptomyces  
RT avermitilis: deducing the ability of producing secondary  
RT metabolites.";  
RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
DR EMBL:BA000030; BAC69721.1; -; Genomic DNA.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 208 AA; 20876 MW; 4CDS5B8C11AF502 CRC64;  
  
Query Match 57.1%; Score 16; DB 2; Length 208;  
Best Local Similarity 16.7%; Pred. No. 2e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 54 CPAASSASAAAC 65  
  
RESULT 32  
Q67VV9_ORYSA
```

```
ID Q67VV9_ORYSA PRELIMINARY; PRT; 210 AA.  
AC Q67VV9;  
DT 25-OCT-2004 (TRMBLrel. 28, Created)  
DT 25-OCT-2004 (TRMBLrel. 28, Last sequence update)  
DT 01-FEB-2005 (TRMBLrel. 29, Last annotation update)  
DE Hypothetical protein OSUNBA0023122.2 (Hypothetical protein  
DE P0530H05.17).  
GN Name=OSUNBA0023122.2; Synonym=P0530H05.17;  
OS Oryza sativa (japonica cultivar-group);  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzoideae; Oryza.  
OC NCBI_Taxid=39947;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC  
RT clone:OSUNBA0023122.";  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC  
RT clone:P0530H05.";  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL:AP004733; BAD37710.1; -; Genomic_DNA.  
DR EMBL:AP003541; BAD37385.1; -; Genomic_DNA.  
DR Gramene; Q67VV9; -;  
KW Hypothetical protein.  
SQ SEQUENCE 210 AA; 21575 MW; C5EB3FFP3B9F4457 CRC64;  
  
Query Match 57.1%; Score 16; DB 2; Length 210;  
Best Local Similarity 16.7%; Pred. No. 2e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 57 CSTSSSSSSRR 68  
  
RESULT 33  
Q5K9J0_CRYNE PRELIMINARY; PRT; 214 AA.  
ID Q5K9J0;  
AC Q5K9J0;  
DT 10-MAY-2005 (TRMBLrel. 30, Created)  
DT 10-MAY-2005 (TRMBLrel. 30, Last sequence update)  
DT 10-MAY-2005 (TRMBLrel. 30, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=CNK01770;  
OS Cryptococcus neoformans var. neoformans JEC21.  
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;  
OC Tremellales; Tremellales; Tremellaceae; Filobasidiella.  
OC NCBI_Taxid=214684;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=JEC21.  
RA Loftus B., Amedeo P., Roncaglia P., Vamathavan J., Uterback T.,  
RA Van Aken S., Frazer C.;  
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).  
RC STRAIN=JEC21;  
RC PubMed=15653466; DOI=10.1126/science.1103773;  
RX Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,  
RA Vamathavan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,  
RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,  
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,  
RA Huang J.C., Jambon G., Jones S.J.M., Koo H.L., Krzywinski M.I.,  
RA Kwon-Chung K.J., Lengele K.B., Maiti R., Maiti M.A., Maiti K.B.,  
RA Mathewson C.A., Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L.,  
RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,  
RA Suh B.B., Tenney A., Uterback T.R., Wicks B.L., Wortman J.R.,  
RA Mye N.H., Kronstad J.W., Lodge J.K., Heltman J., Davis R.W.,
```

RA Fraser C.M., Hyman R.W.;
RT "The genome of the basidiomycetous yeast and human pathogen
RT Cryptococcus neoformans.";
RT Science 307:1321-1324 (2005).
DR EMBL: AE017351; AAW46337.1; -; Genomic_DNA.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0005732; C:small nuclear ribonucleoprotein complex; IEA.
DR GO: GO:0006397; P:mRNA processing; IEA.
DR InterPro: IPR001163; LSM_snrnp_core.
DR Pfam: PF01423; LSM; 1.
DR SMART; SM00651; Sm; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 214 AA; 23828 MW; 9DEFCED4844BCE42 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 214;
Best Local Similarity 16.7%; Pred. No. 2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 9 CVASSTSSVASC 20

RESULT 34
O5SK09 CRYNE PRELIMINARY; PRT; 215 AA.
ID O5SK09;
AC O5SK09;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CNBK1770;
OS Cryptococcus neoformans var. neoformans B-3501A.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OK NCBI_TaxID=283643;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B-3501A;
RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
RA Wicks B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing.";
RT Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAEY0100052; EAL18156.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 215 AA; 24014 MW; 1CEDEFCED4844BCE CRC64;

Query Match 57.1%; Score 16; DB 2; Length 215;
Best Local Similarity 16.7%; Pred. No. 2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 9 CVASSTSSVASC 20

RESULT 35
O5K919 CRYNE PRELIMINARY; PRT; 215 AA.
ID O5K919;
AC O5K919;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CNK01770;
OS Cryptococcus neoformans var. neoformans JEC21.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OK NCBI_TaxID=214684;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=JEC21;
RA Loftus B., Amedeo P., Roncaglia P., Vamathevan J., Uterback T.,
RA Van Aken S., Fraser C.; to the EMBL/GenBank/DBJ databases.
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=JEC21;
RX PubMed=15653466; DOI=10.1126/science.1103773;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA Bosdet I.B., Brent M.R., Chin R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Janbon G., Jones S.J.M., Koo H.L., Krzywinski M.I.,
RA Kwon-Chung K.J., Lengeler K.B., Maiti R., Marra M.A., Marra R.E.,
RA Mathewson C.A., Mitchell T.G., Petrea M., Riggs F.R., Salzberg S.L.,
RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
RA Suh B.B., Tenney A., Uterback T.R., Wicks B.L., Wortman J.R.,
RA Wye N.H., Kronstad J.W., Lodge J.K., Heltman J., Davis R.W.,
RA Fraser C.M., Hyman R.W.;
RT "The genome of the basidiomycetous yeast and human pathogen
RT Cryptococcus neoformans.";
RL Science 307:1321-1324 (2005).
DR EMBL: AE017351; AAW46338.1; -; Genomic_DNA.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0005732; C:small nuclear ribonucleoprotein complex; IEA.
DR GO: GO:0006397; P:mRNA processing; IEA.
DR InterPro: IPR001163; LSM_snrnp_core.
DR Pfam: PF01423; LSM; 1.
DR SMART; SM00651; Sm; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 215 AA; 24014 MW; 1CEDEFCED4844BCE CRC64;

Query Match 57.1%; Score 16; DB 2; Length 215;
Best Local Similarity 16.7%; Pred. No. 2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 9 CVASSTSSVASC 20

RESULT 36
O4QSJ9 GLECA PRELIMINARY; PRT; 227 AA.
ID O4QSJ9;
AC O4QSJ9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative non-reducing polyketide synthase (Fragment).
OS Pertusaria coronata.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Lecanoromycetes;
OC Pertusariales; Pertusariaceae; Pertusaria.
OK NCBI_TaxID=232756;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15927215; DOI=10.1016/j.phytochem.2005.04.014;
RA Schmitt I., Martin M.P., Kautz S., Lumbach H.T.;
RT "Diversity of non-reducing polyketide synthase genes in the
RT Pertusariales (lichenized Ascomycota): A phylogenetic perspective.";
RT Phytochemistry 66:1241-1253 (2005).
DR EMBL: AY918748; AAY00086.1; -; Genomic_DNA.
DR InterPro: IPR000794; Ketoacyl_synth.
DR Pfam: PF00109; ketoacyl-synth; 1.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 1.
KW Transferase.
FT NON_TER 227
FT NON_TER 227
SQ SEQUENCE 227 AA; 24445 MW; 4D57BB346B626A8C CRC64;

Query Match 57.1%; Score 16; DB 2; Length 227;
Best Local Similarity 16.7%; Pred. No. 2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12
 Db 95 CSSSFAAISTAC 106

RESULT 37
 Q7JNB5_GIALA PRELIMINARY; PRT; 230 AA.
 AC Q7JNB5;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Variant-specific surface protein type 4 tsapl/tsa417-like
 DE (Fragment).
 GN Name=vsp417-4;
 OS Eukaryota; Diplomonadida; Hexamitidae; Giardia.
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.
 NCBI_TaxID=5741;
 RX MEDLINE=99229863; PubMed=10215024; DOI=10.1016/S0166-6851(98)00183-2;
 RA E. P. L., Darby J. M., Mayrhofer G.;
 RT "A new locus (vsp417-4) belonging to the tsa417-like subfamily of
 RT variant-specific surface protein genes in Giardia intestinalis."
 RL Mol. Biochem. Parasitol. 93:55-68(1999).
 DR EMBL; U89153; AAD05041.1; -; Genomic_DNA.
 DR InterPro; IPR006212; Furin repeat.
 DR InterPro; IPR005127; Giardia_VSP.
 DR InterPro; IPR006210; IEGF.
 DR Pfam; PF03302; VSP; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00261; FU; 1.
 FT NON_TER 1
 FT NON_TER 230
 FT NON_TER 230
 SQ SEQUENCE 230 AA; 23336 MW; 45E22ECD1C0084F1 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 230;
 Best Local Similarity 16.7%; Pred. No. 2e-07;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12
 Db 22 CSSTVRAAATC 33

RESULT 38
 Q8LD53_ARATH PRELIMINARY; PRT; 234 AA.
 AC Q8LD53;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE BHLH transcription factor, putative.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC Rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RX MEDLINE=22088475; PubMed=12093376;
 RA Haas B. J., Volfovsky N., Town C. D., Troukhan M., Alexandrov N.,
 RA Feldmann K. A., Flavell R. B., White O., Salzberg S. L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 RT annotation";
 RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
 RP NUCLEOTIDE SEQUENCE.
 RP NUCLEOTIDE SEQUENCE.
 RP Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (BHLH) domain.

DR EMBL; AY086197; AAM64276.1; -; mRNA.
 DR HSSP; P36956; IAM9.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0030528; F:transcription regulator activity; IEA.
 DR GO; GO:0045449; P:regulation of transcription; IEA.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS50889; HLH; 1.
 SQ SEQUENCE 234 AA; 25470 MW; 9EBEA7D3E44F21D5 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 234;
 Best Local Similarity 16.7%; Pred. No. 2e-07;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12
 Db 67 CESSSATSSKAC 78

RESULT 39
 Q9FH37_ARATH PRELIMINARY; PRT; 234 AA.
 AC Q9FH37;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-FEB-2003 (TrEMBLrel. 29, Last annotation update)
 DE Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:K5F14
 DE (Hypothetical protein At5g54680; K5F14.2) (Putative bHLH transcription
 DE factor) (Hypothetical protein At5g54680).
 GN Name=At5g54680/K5F14.2; Synonyms=At5g54680;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC Rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RX MEDLINE=20181125; PubMed=10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
 RT clones.";
 RL DNA Res. 7:31-63(2000).
 RP NUCLEOTIDE SEQUENCE.
 RP Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C. J., Bower L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M. K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S. X., Narusaka M., Pham P. K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinzaki K.,
 RA Ecker J., Theologis A., Davis R. W.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 RP NUCLEOTIDE SEQUENCE.
 RP MEDLINE=225398051; PubMed=12679534; DOI=10.1093/molbev/msg088;
 RA Heim W. A., Jakoby M., Weber M., Martin C., Weisshar B., Bailey P. C.;
 RT "The basic helix-loop-helix transcription factor family in plants: a
 RT genome-wide study of protein structure and functional diversity.";
 RL Mol. Biol. Evol. 20:735-747(2003).
 RP NUCLEOTIDE SEQUENCE.
 RP Jakoby M. J., Heim M. A., Bailey P., Martin C., Weisshar B.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RP NUCLEOTIDE SEQUENCE.
 RP Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
 RA Palm C. J., Bower L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M. K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S. X., Narusaka M., Pham P. K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinzaki K.,
 RA Ecker J., Theologis A., Davis R. W.;

RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 DR EMBL; AB022214; BAB09934.1; -; Genomic_DNA.
 DR EMBL; AY054585; AAK96776.1; -; mRNA.
 DR EMBL; AF488629; AAM10964.1; -; mRNA.
 DR EMBL; BT002189; AAM72200.1; -; mRNA.
 DR HSSP; P36956; 1AM9.
 DR GO; GO:0045449; P:regulation of transcription; TAS.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS00888; HLH; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 234 AA; 25488 MW; 987BA114244F21D1 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 234;
 Best Local Similarity 16.7%; Pred. No. 2e-07;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12
 DB 67 CESSATSSKAC 78

RESULT 40
 016421 CAEEL PRELIMINARY; PRT; 244 AA.
 ID 016421;
 AC 016421;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DE 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Hypothetical protein T05B4.12.
 GN ORFNames=T05B4.12;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RX [1]_TaxID=6239;
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Bristol N2;
 RC MEDLINE=99069613; PubMed=9851916;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; AF016445; AAC69052.1; -; Genomic_DNA.
 DR PIR; T31838; T31838.
 DR Ensembl; T05B4.12; Caenorhabditis elegans.
 DR Wormbase; WBGene00020243; T05B4.12.
 DR WormPep; T05B4.12; CEI3215.
 DR InterPro; IPR003582; ShKT.
 DR Pfam; PF01549; ShKT; 3.
 DR SMART; SM00254; ShKT; 4.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 244 AA; 25347 MW; 36362492AC6B1C9 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 244;
 Best Local Similarity 16.7%; Pred. No. 2.1e-07;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12
 DB 203 CTSYADSSSTSC 214

RESULT 41
 0966B6 CAEEL PRELIMINARY; PRT; 247 AA.
 ID 0966B6;
 AC 0966B6;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
 DE Hypothetical protein.

GN ORFNames=Y46H3D.8;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RX [1]_TaxID=6239;
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Bristol N2;
 RC MEDLINE=99069613; PubMed=9851916;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; AC006777; AAK72311.1; -; Genomic_DNA.
 DR Ensembl; Y46H3D.8; Caenorhabditis elegans.
 DR InterPro; IPR003582; ShKT.
 DR Pfam; PF01549; ShKT; 3.
 DR SMART; SM00254; ShKT; 4.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 247 AA; 25215 MW; FF3CD3BF00DE5A7 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 247;
 Best Local Similarity 16.7%; Pred. No. 2.1e-07;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12
 DB 206 CTSYADSSSAC 217

RESULT 42
 016424 CAEEL PRELIMINARY; PRT; 248 AA.
 ID 016424;
 AC 016424;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DE 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Hypothetical protein T05B4.3.
 GN ORFNames=T05B4.3;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RX [1]_TaxID=6239;
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Bristol N2;
 RC MEDLINE=99069613; PubMed=9851916;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; AF016445; AAC69054.1; -; Genomic_DNA.
 DR PIR; T31841; T31841.
 DR Ensembl; T05B4.3; Caenorhabditis elegans.
 DR Wormbase; WBGene00020237; T05B4.3.
 DR WormPep; T05B4.3; CEI3197.
 DR InterPro; IPR003582; ShKT.
 DR Pfam; PF01549; ShKT; 3.
 DR SMART; SM00254; ShKT; 4.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 248 AA; 25659 MW; 79B281AFBBD3AFC CRC64;

Query Match 57.1%; Score 16; DB 2; Length 248;
 Best Local Similarity 16.7%; Pred. No. 2.1e-07;
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12
 DB 207 CTSFADSSSSC 218

RESULT 43
 09XV90 CAEEL

ID Q9XV90_CABEL PRELIMINARY; PRT; 257 AA.
AC Q9XV90;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein Fl6H6.3
GN Name=Fl6H6.3; ORFName=Fl6H6.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Briscot N2;
RA Matthews L.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Briscot N2;
RX MEDLINE=99069613; PubMed=9851916; DOI=10.1126/science.282.5396.2012;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: Z81506; CAB04129.1; -; Genomic_DNA.
DR PIR: T21029; T21029.
DR Ensemble; Fl6H6.3; Caenorhabditis elegans.
DR Wormbase; WBGene0000893; Fl6H6.3.
DR WormPep; Fl6H6.3; CE18599.
DR InterPro; IPR003582; ShKT.
DR Pfam; PF01549; ShTK; 3.
DR SMART; SM00254; ShTK; 4.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 257 AA; 27646 MW; CA204FLA6C00EB87 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 257;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 195 CTNATTTTQASC 206

RESULT 44
Q8CAZ5_MOUSE PRELIMINARY; PRT; 262 AA.
AC Q8CAZ5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Mus musculus 6 days neonate skin cDNA, RIKEN full-length enriched
library, clone:A030004L08 product:hypothetical protein, full insert
sequence. (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RL "High-efficiency full-length cDNA cloning."
RT Meth. Enzymol. 303:19-44(1999).
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiyama H.,
RA Kuenl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Bash G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.U., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofman M., Hume D.A., Kamlya M., Lee N.H.,
RA Lyons P., Marchionni L., Maehima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Tovo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszew-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,
RA Hayashizaki Y.;
RL "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RG The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/9r.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugihara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/9r.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Saeki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Katsunai T., Taahiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wachihi M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Tanaka S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-Format
sequencing pipeline with 384 multiplexed sequencer."
RL Genome Res. 10:1757-1771(2000).
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirose T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ono M., Onisato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Taya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK037168; BAC29730.1; -; mRNA.
DR Ensemble; ENSMUSG0000050239; Mus musculus.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 262 AA; 29111 MW; 8D441316CD4432E8 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 262;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12
Db 139 CTTATRSASKAC 150

RESULT 45

ID Q7PRO7 ANOGA PRELIMINARY; PRT; 269 AA.
AC Q7PRO7;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DE ENSANGP0000001657 (Fragment).
GN ORFNames=ENSANG0000001387;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OX Anophelinae; Anopheles.
[1] NCBI_TaxID=180454;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA01008847; FA06779.3; -; Genomic_DNA.
FT NON_TER 1 1
FT NON_TER 269 269
SQ SEQUENCE 269 AA; 24228 MW; 25BBFF71FD71F1F2 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 269;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12
Db 114 CASSCSTSSC 125

RESULT 46

ID Q6ZUB7 HUMAN PRELIMINARY; PRT; 278 AA.
AC Q6ZUB7;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE Hypothetical protein FLN43840.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo
OX NCBI_TaxID=9606;
[1]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Testis;
RC Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Ohshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K.,
RA Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK125828; BAC6309.1; -; mRNA.

SQ SEQUENCE 278 AA; 27768 MW; B0C6757FC1DAE2C0 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 278;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12
Db 35 CTAAGSSSLSTC 46

RESULT 47

ID O8PPE6 XANAC PRELIMINARY; PRT; 285 AA.
AC O8PPE6;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
GN Endonuclease.
OS OrderdiocunNames=XAC4037;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Fairah C.S., Furian L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F., Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Gamariotte G., Camavan F., Cardoso J., Chambergo F.,
RA Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorri H., Faria J.B., Ferreira A.U.B., Ferreira R.C.C.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sene J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RU Nature 417:459-463(2002).
DR EMBL; AE012052; AAM38872.1; -; Genomic_DNA.
DR HSSP; P13717; 1G8T.
DR GO; GO:0004519; P:endonuclease activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001604; Endonuclease.
DR Pfam; PF01223; Endonuclease_NS; 1.
DR SMART; SM00477; NUC; 1.
DR PROSITE; PS01070; NUCLEASE_NON_SPEC; UNKNOWN_1.
KW Complete proteome; Endonuclease.

SQ SEQUENCE 285 AA; 30358 MW; 5581A0B49A2057BB CRC64;

Query Match 57.1%; Score 16; DB 2; Length 285;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12
Db 51 CTTVAATTRC 62

RESULT 48

ID Q6ZOO1 HUMAN PRELIMINARY; PRT; 286 AA.
AC Q6ZOO1;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

```

DE Hypochemical protein FLJ46328.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RA Niinomiya K., Wagatsuna M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa K., Yamazaki M., Sugiyama T., Irie R.,
RA Katsuba N., Sato K., Tanikawa M., Yamazaki M., Ishii S., Yamamoto J., Isono Y.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hiro Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanohori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugeno S., Nagahari K., Maunho Y., Nagai K., Isogai T.,
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF128833; BACB7633.1; -; mRNA.
SQ SEQUENCE 286 AA; 30415 MW; B6CEAPFA76737B8C CRC64;

Query Match 57.1%; Score 16; DB 2; Length 286;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 231 CTAAGSSSLSTC 242

RESULT 49
Q9F934 BIFAD PRELIMINARY; PRT; 287 AA.
AC Q9F934;
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DE 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Sugar binding protein Sbp.
GN Name=sbp;
OS Bifidobacterium adolescentis.
OC Bacteria; Actinobacteria; Actinobacteriales; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=1680;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=INT57;
RA Park M.S., Yoon H.J., Ji G.B.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF213175; AAC31596.1; -; Genomic_DNA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR006059; SBP_bac_1.
DR Pfam; PF01547; SBP_bac_1; 1.
SQ SEQUENCE 287 AA; 32054 MW; 11B9382AE39B86DB CRC64;

Query Match 57.1%; Score 16; DB 2; Length 287;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 48 CAAIMASAAAC 59

RESULT 50
Q89DV5 BRAJA PRELIMINARY; PRT; 297 AA.
AC Q89DV5;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE B117302 protein.
GN OrderedLocNames=B117302;

```

```

OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=2248498; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamideawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Ideawa K., Iritsugu M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000040; BACS2567.1; -; Genomic_DNA.
SQ SEQUENCE 297 AA; 32798 MW; E8B4D21866645601 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 297;
Best Local Similarity 16.7%; Pred. No. 2.2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 250 CSATHSDDAAC 261

```

Search completed: January 4, 2006, 16:09:56
 Job time : 79.3652 secs